

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 September 2001 (07.09.2001)

PCT

(10) International Publication Number  
**WO 01/64920 A2**

(51) International Patent Classification<sup>7</sup>: **C12N 15/62**,  
C07K 14/22, 19/00

Siena (IT). **PIZZA, Mariagrazia** [IT/IT]; Chiron SpA,  
Via Fiorentina, 1, I-53100 Siena (IT).

(21) International Application Number: PCT/IB01/00420

(74) Agents: **HALLYBONE, Huw, George et al.**; Carpinaels  
& Ransford, 43 Bloomsbury Square, London WC1A 2RA  
(GB).

(22) International Filing Date: 28 February 2001 (28.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
0004695.3 28 February 2000 (28.02.2000) GB  
0027675.8 13 November 2000 (13.11.2000) GB

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,  
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(71) Applicant (*for all designated States except US*): **CHIRON  
SPA** [IT/IT]; Via Fiorentina, 1, I-53100 Siena (IT).

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

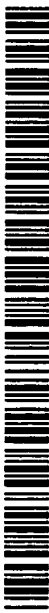
(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ARICO, Maria**,  
**Beatrice** [IT/IT]; Chiron SpA, Via Fiorentina 1, I-53100  
Siena (IT). **COMANDUCCI, Maurizio** [IT/IT]; Chiron  
SpA, Via Fiorentina 1, I-53100 Siena (IT). **GALEOTTI**,  
**Cesira** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100  
Siena (IT). **MASIGNANI, Vega** [IT/IT]; Chiron SpA, Via  
Fiorentina, 1, I-53100 Siena (IT). **GIULIANI, Marzia**,  
**Monica** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100

Published:

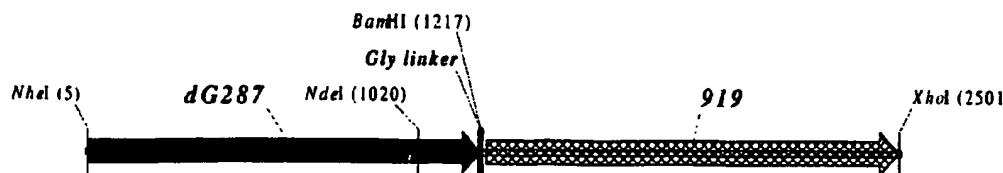
— without international search report and to be republished  
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/64920 A2

(54) Title: HYBRID EXPRESSION OF NEISSERIAL PROTEINS



(57) Abstract: Two or more Neisserial proteins (e.g. A and B) are expressed as a single hybrid protein which can be represented simply by the formula NH<sub>2</sub>-A-B-COOH.

## HYBRID EXPRESSION OF NEISSERIAL PROTEINS

All documents cited herein are incorporated by reference in their entirety.

### TECHNICAL FIELD

This invention is in the field of protein expression. In particular, it relates to the heterologous  
5 expression of proteins from *Neisseria* (e.g. *N.gonorrhoeae* or, preferably, *N.meningitidis*).

### BACKGROUND ART

International patent applications WO99/24578, WO99/36544, WO99/57280 and  
WO00/22430 disclose proteins from *Neisseria meningitidis* and *Neisseria gonorrhoeae*.  
These proteins are typically described as being expressed in *E.coli* (i.e. heterologous  
10 expression) as either N-terminal GST-fusions or C-terminal His-tag fusions, although other  
expression systems, including expression in native *Neisseria*, are also disclosed.

It is an object of the present invention to provide alternative and improved approaches for  
the heterologous expression of these proteins. These approaches will typically affect the  
level of expression, the ease of purification, the cellular localisation of expression, and/or the  
15 immunological properties of the expressed protein.

### DISCLOSURE OF THE INVENTION

In accordance with the invention, two or more (e.g. 3, 4, 5, 6 or more) proteins of the  
invention are expressed as a single hybrid protein. It is preferred that no non-Neisserial  
fusion partner (e.g. GST or poly-His) is used.

20 This offers two advantages. Firstly, a protein that may be unstable or poorly expressed on its  
own can be assisted by adding a suitable hybrid partner that overcomes the problem.  
Secondly, commercial manufacture is simplified – only one expression and purification need  
be employed in order to produce two separately-useful proteins.

Thus the invention provides a method for the simultaneous heterologous expression of two  
25 or more proteins of the invention, in which said two or more proteins of the invention are  
fused (i.e. they are translated as a single polypeptide chain).

The method will typically involve the steps of: obtaining a first nucleic acid encoding a first  
protein of the invention; obtaining a second nucleic acid encoding a second protein of the

invention; ligating the first and second nucleic acids. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

Where just two proteins are joined, the hybrid protein can be represented simply by the formula  $\text{NH}_2\text{-A-B-COOH}$ . A and B can each be selected from any Neisserial proteins, and in particular those represented by SEQ#s 1-4326. The method is well suited to the expression of proteins orf1, orf4, orf25, orf40, Orf46/46.1, orf83, 233, 287, 292L, 564, 687, 741, 907, 919, 953, 961 and 983.

The 42 hybrids indicated by 'X' in the following table of form  $\text{NH}_2\text{-A-B-COOH}$  are preferred:

A → B	ORF46.1	287	741	919	953	961	983
ORF46.1		X	X	X	X	X	X
287	X		X	X	X	X	X
741	X	X		X	X	X	X
919	X	X	X		X	X	X
953	X	X	X	X		X	X
961	X	X	X	X	X		X
983	X	X	X	X	X	X	

Preferred proteins to be expressed as hybrids are thus ORF46.1, 287, 741, 919, 953, 961 and 983. These may be used in their essentially full-length form, or poly-glycine deletions ( $\Delta\text{G}$ ) forms may be used (e.g.  $\Delta\text{G-287}$ ,  $\Delta\text{GTbp2}$ ,  $\Delta\text{G741}$ ,  $\Delta\text{G983}$  etc.), or truncated forms may be used (e.g.  $\Delta\text{1-287}$ ,  $\Delta\text{2-287}$  etc.), or domain-deleted versions may be used (e.g. 287B, 287C, 287BC, ORF46<sub>1-433</sub>, ORF46<sub>433-608</sub>, ORF46, 961c etc.) and so on.

Particularly preferred are: (a) a hybrid protein comprising 919 and 287; (b) a hybrid protein comprising 953 and 287; (c) a hybrid protein comprising 287 and ORF46.1; (d) a hybrid protein comprising ORF1 and ORF46.1; (e) a hybrid protein comprising 919 and ORF46.1; (f) a hybrid protein comprising ORF46.1 and 919; (g) a hybrid protein comprising ORF46.1, 287 and 919; (h) a hybrid protein comprising 919 and 519; and (i) a hybrid protein comprising ORF97 and 225.

Further embodiments are shown in the drawings and include  $\Delta\text{G287-919}$ ,  $\Delta\text{G287-953}$ ,  $\Delta\text{G287-961}$ ,  $\Delta\text{G983-ORF46.1}$ ,  $\Delta\text{G983-741}$ ,  $\Delta\text{G983-961}$ ,  $\Delta\text{G983-961C}$ ,  $\Delta\text{G741-961}$ ,  $\Delta\text{G741-961C}$ ,  $\Delta\text{G741-983}$ ,  $\Delta\text{G741-ORF46.1}$ , ORF46.1-741, ORF46.1-961, ORF46.1-961C,

961-ORF46.1, 961-741, 961-983, 961C-ORF46.1, 961C-741, 961C-983, 961CL-ORF46.1, 961CL-741, and 961CL-983.

Where 287 is used, it is preferably at the C-terminal end of a hybrid; if it is to be used at the N-terminus, it is preferred to use a  $\Delta G$  form of 287 is used (*e.g.* as the N-terminus of a hybrid with ORF46.1, 919, 953 or 961).

Where 287 is used, this is preferably from strain 2996 or from strain 394/98.

Where 961 is used, this is preferably at the N-terminus. Domain forms of 961 may be used.

Alignments of polymorphic forms of ORF46, 287, 919 and 953 are disclosed in WO00/66741. Any of these polymorphs can be used according to the present invention.

10 Preferably, the constituent proteins (A and B) in a hybrid protein according to the invention will be from the same strain.

The fused proteins in the hybrid may be joined directly, or may be joined via a linker peptide *e.g.* via a poly-glycine linker (*i.e.*  $G_n$  where  $n = 3, 4, 5, 6, 7, 8, 9, 10$  or more) or via a short peptide sequence which facilitates cloning. It is evidently preferred not to join a  $\Delta G$  protein to the C-terminus of a poly-glycine linker.

The fused proteins may lack native leader peptides or may include the leader peptide sequence of the N-terminal fusion partner.

### **Host**

It is preferred to utilise a heterologous host. The heterologous host may be prokaryotic or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonella typhimurium*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Neisseria lactamica*, *Neisseria cinerea*, *Mycobacteria* (*e.g.* *M.tuberculosis*), yeast *etc.*

### **Vectors, hosts etc.**

25 As well as the methods described above, the invention provides (a) nucleic acid and vectors useful in these methods (b) host cells containing said vectors (c) proteins expressed or expressable by the methods (d) compositions comprising these proteins, which may be suitable as vaccines, for instance, or as diagnostic reagents, or as immunogenic compositions (e) these compositions for use as medicaments (*e.g.* as vaccines) or as diagnostic reagents (f)

the use of these compositions in the manufacture of (1) a medicament for treating or preventing infection due to Neisserial bacteria (2) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria, and/or (3) a reagent which can raise antibodies against Neisserial bacteria and (g) a method of treating a patient, comprising administering to the patient a therapeutically effective amount of these compositions.

### ***Sequences***

The invention also provides a protein or a nucleic acid having any of the sequences set out in the following examples. It also provides proteins and nucleic acid having sequence identity to these. As described above, the degree of 'sequence identity' is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more).

### ***Nomenclature herein***

The 2166 protein sequences disclosed in WO99/24578, WO99/36544 and WO99/57280 are referred to herein by the following SEQ# numbers:

<b>Application</b>	<b>Protein sequences</b>	<b>SEQ# herein</b>
WO99/24578	Even SEQ IDs 2-892	SEQ#s 1-446
WO99/36544	Even SEQ IDs 2-90	SEQ#s 447-491
WO99/57280	Even SEQ IDs 2-3020	SEQ#s 492-2001
	Even SEQ IDs 3040-3114	SEQ#s 2002-2039
	SEQ IDs 3115-3241	SEQ#s 2040-2166

In addition to this SEQ# numbering, the naming conventions used in WO99/24578, WO99/36544 and WO99/57280 are also used (*e.g.* 'ORF4', 'ORF40', 'ORF40-1' *etc.* as used in WO99/24578 and WO99/36544; 'm919', 'g919' and 'a919' *etc.* as used in WO99/57280).

The 2160 proteins NMB0001 to NMB2160 from Tettelin *et al.* [*Science* (2000) 287:1809-1815] are referred to herein as SEQ#s 2167-4326 [see also WO00/66791].

The term 'protein of the invention' as used herein refers to a protein comprising:

- (a) one of sequences SEQ#s 1-4326; or
- (b) a sequence having sequence identity to one of SEQ#s 1-4326; or
- (c) a fragment of one of SEQ#s 1-4326.

The degree of 'sequence identity' referred to in (b) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). This includes mutants and allelic variants [e.g. see WO00/66741]. Identity is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters *gap open penalty=12* and *gap extension penalty=1*. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence.

The 'fragment' referred to in (c) should comprise at least *n* consecutive amino acids from one of SEQ#s 1-4326 and, depending on the particular sequence, *n* is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragment comprises an epitope from one of SEQ#s 1-4326. Preferred fragments are those disclosed in WO00/71574 and WO01/04316.

Preferred proteins of the invention are found in *N.meningitidis* serogroup B.

Preferred proteins for use according to the invention are those of serogroup B *N.meningitidis* strain 2996 or strain 394/98 (a New Zealand strain). Unless otherwise stated, proteins mentioned herein are from *N.meningitidis* strain 2996. It will be appreciated, however, that the invention is not in general limited by strain. References to a particular protein (e.g. '287', '919' etc.) may be taken to include that protein from any strain.

It will be appreciated that references to "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA) etc.

## BRIEF DESCRIPTION OF DRAWINGS

Figures 1 to 26 show hybrid proteins according to the invention.

## MODES FOR CARRYING OUT THE INVENTION

### 25 *Example 1 – hybrids of ORF46*

The complete ORF46 protein from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

30                   1    LGISRKISLI   LSILAVCLPM   HAHASDLAND   SFIRQVLDRO   HFEPDGKYHL  
                   51    FGSRGELAER   SGHIGLGKIQ   SHQLGNLMIQ   QAAIKGNIGY   IVRFSDHGHE  
                   101   VHSPFDNHAS   HSDSDEAGSP   VDGFSLYRIH   WDGYEHHPAD   GYDGPQGGGY

151 PAPKGARDIY SYDIKGVAQN IRLNLTNRS TGQRLADRFH NAGSMLTQGV  
 201 GDGFKRATRY SPELDRSGNA AEAFTGTADI VKNIIGAAGE IVGAGDAVQG  
 251 ISEGSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA AIRDWAVQNP  
 301 NAAQGIEAVS NIFMAAIPK GIGAVRGKYG LGGITAHPIK RSQMGAIALP  
 351 KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LEQRYGKENI TSSTVPPSNG  
 401 KNVKLADQRH PKTGVPPFDGK GFPNFEKHVK YDTKLDIQEL SGGGIPKAKP  
 451 VSDAKPRWEV DRKLNKLTR EQVEKNVQEI RGNKNKSNFS QHAQLEREIN  
 501 KLKSADEINF ADGMGKFTDS MNDKAFSRLV KSVKENGFTN PVVEYVEING  
 551 KAYIVRGNNR VFAAEYLGR I HELKFKKVDF FVPNTSWKNP TDVLNESGNV  
 601 KRPRYRSK\*

The leader peptide is underlined.

The sequences of ORF46 from other strains can be found in WO00/66741.

ORF46 has been fused at its C-terminus and N-terminus with 287, 919, and ORF1. The hybrid proteins were generally insoluble, but gave some good ELISA and bactericidal results (against the homologous 2996 strain):

Protein	ELISA	Bactericidal Ab
Orf1-Orf46.1-His	850	256
919-Orf46.1-His	12900	512
919-287-Orf46-His	n.d.	n.d.
Orf46.1-287His	150	8192
Orf46.1-919His	2800	2048
Orf46.1-287-919His	3200	16384

For comparison, 'triple' hybrids of ORF46.1, 287 (either as a GST fusion, or in  $\Delta$ G287 form) and 919 were constructed and tested against various strains (including the homologous 2996 strain) *versus* a simple mixture of the three antigens. FCA was used as adjuvant:

	2996	BZ232	MC58	NGH38	F6124	BZ133
Mixture	8192	256	512	1024	>2048	>2048
ORF46.1-287-919his	16384	256	4096	8192	8192	8192
$\Delta$ G287-919-ORF46.1his	8192	64	4096	8192	8192	16384
$\Delta$ G287-ORF46.1-919his	4096	128	256	8192	512	1024

Again, the hybrids show equivalent or superior immunological activity.

Hybrids of two proteins (strain 2996) were compared to the individual proteins against various heterologous strains:

	1000	MC58	F6124 (MenA)
ORF46.1-His	<4	4096	<4
ORF1-His	8	256	128
ORF1—ORF46.1-His	1024	512	1024

Again, the hybrid shows equivalent or superior immunological activity.

**Example 2 – hybrids of  $\Delta$ G287**

The deletion of the (Gly)<sub>6</sub> sequence in 287 was found to have a dramatic effect on protein expression. The protein lacking the N-terminal amino acids up to GGGGGG is called ' $\Delta$ G287'. In strain MC58, its basic sequence (leader peptide underlined) is:

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40

```

SPDVKS ADTLSKPAAP VVSEKETEAQ EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT
GNNGAVTADN PKNEDEVAQN DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP
ANQPDMANAA DGMQGGDDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
NPGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS DADKISNYKK
DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRS ARSRRLPAE MPLIPVNQAD
TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY
NGEVLHFPTE NGRPYPTRGR FAAKVDFGSK SVDGIIDSGD DLHMGTKQFK AAIDGNGFKG
TWTENGSGDV SGKFYGPAGE EVAGKYSYRP TDAEKGFGV FAGKKEQD*
```

15  $\Delta$ G287, with or without His-tag (' $\Delta$ G287-His' and ' $\Delta$ G287K', respectively), are expressed at very good levels in comparison with the '287-His' or '287<sub>untagged</sub>'.

On the basis of gene variability data, variants of  $\Delta$ G287-His were expressed in *E.coli* from a number of MenB strains, in particular from strains 2996, MC58, 1000, and BZ232. The results were also good – each of these gave high ELISA titres and also serum bactericidal titres of >8192.  $\Delta$ G287K, expressed from pET-24b, gave excellent titres in ELISA and the serum bactericidal assay.

Deletion of poly-Gly sequences is also applicable to Tbp2 (NMB0460), 741 (NMB 1870) and 983 (NMB1969). When cloned in pET vector and expressed in *E.coli* without the sequence coding for their leader peptides and without poly-Gly (*i.e.* as " $\Delta$ G forms"), the same effect was seen – expression was good in the clones carrying the deletion of the poly-glycine stretch, and poor or absent if the glycines were present in the expressed protein.

$\Delta$ G287 was fused directly in-frame upstream of 919, 953, 961 (sequences shown below) and ORF46.1:

30  
 35  
 40

```

 $\Delta$ G287-919
ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
GATATGCCGCAAAATTCGCCGAATCCGCAAAATCAAACAGGGAACAACCAACCCGCCGATTCTTCAGATTCCGCCCCC
GCGTCAAACCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAGGGTTGATTGGCTAATGGCGTTTGTATTGATGGG
CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAAATGGTGATAATTTATTGGATGAAGAAGCACCG
TCAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAAATTT
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACATAACAAATATGTATCATCATTTATAAAGACAAGTCCGCTTCA
TCTTCATCTGCGCGATTGAGGCGTTCTGCACGGTCGAGGAGGTCGCTTCTGCGGAGATGCCGCTAATCCCGTCAAT
CAGGCGGATACGCTGATTGTGATGGGGAAGCGGTCAGCCTGACGGGGCATTCCGGCAATATCTTCGCGCCCCGAAGGG
AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCCGCGGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA
AAAGGCGAAATGCTTGTGCGCACGGCCGTGTACAACGGCGAAGTGCTGCATTTTCATACGGAACCGCCGTCCGTAC
CCGACTAGAGGCAGGTTTGC CGCAAAAGTCGATTTCGGCAGCAAAATCTGTGGACGGCATTATCGACAGCGCGGATGAT
TTCATATGGGTACGCAAAATTCAAAGCCGCCATCGATGGAACGGCTTTAAGGGGACTTGGACGGAATAATGGCGGC
GGGATGTTTCGGAAGGTTTACGGCCCCGCCGCGGAGGAAGTGCGGGGAAAATACAGCTATCGCCCGACAGATGCG
```



5 GAAAAGGGCGGATTCGGCGTGTTCGCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGC  
ATCCAAACCTTTCCGCAACCCGACACATCCGTCATCAACGGCCCCGACCGCGCGTCCGGCATCCCCGACCCCGCCGGA  
ACGACGGTCCGGCGCGCGGGCGGCGTCTATACCGTGTACCGCACCTGTCCCTGCCCACTGGGGCGGCGAGGATTTTC  
GCCAAAAGCCTGCAATCCTTCCGCTCGGCTGCGCCAATTTGAAAAACCGCAAGGCTGGCAGGATGTGTGCGCCCAA  
GCCTTTCAAACCCCGTCCATTCTTTTCAGGCAAAACAGTTTGTGAACGCTATTTACGCGCGTGGCAGGTTGCAGGC  
AACGGAAGCCTTGGCGGTACGGTTACCGGCTATTACGAGCCGGTGTGAAGGGCGACAGCGCGGACGGCACAAGCC  
CGCTTCCCGATTTACGGTATTTCCGACGATTTTATCTCCGTCCCCCTGCTGCGGTTTGGCGGAGCGGAAAAGCCCTT  
GTCCGCATCAGGCAGACGGGAAAAACAGCGGCACAATCGACAATACCGGCGGCACACATACCGCCGACCTCTCCCGA  
TTCCCCATCACCGCGCGCACAAACGGCAATCAAAGGCAGGTTTGAAGGAAGCCGCTTCTCCCCCTACCACACGCGCAAC  
CAAATCAACGGCGCGCGCTTGACGGCAAAGCCCCGATACTCGGTTACGCCGAAGACCCCGTTCGAACCTTTTTTTTATG  
CACATCCAAGGCTCGGGCGCTGTGAAAACCCCGTCCGGCAAATACATCCGCATCGGCTATGCCGACAAAAACGAACAT  
CCCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGCTACCTCAAGCTCGGGCAGACCTCGATGCAGGGCATCAA  
GCCTATATGCGGCAAAATCCGCAACGCCTCGCCGAAGTTTGGGTCAAACCCCGAGCTATATCTTTTCCGCGAGCTT  
15 GCGGAAGCAGCAATGACGGTCCCGTCCGGCGCACTGGGCACGCCGTTGATGGGGGAATATGCCGCGCAGTCGACCCG  
CACTACATTACCTTGGGCGCGCCCTTATTTGTCCGACCCGCCATCCGGTTACCCGCAAAGCCCTCAACCGCCTGATT  
ATGGCGCAGGATACCGCGCAGCGGATTAAAGGCGCGGTGCGCGTGGATTATTTTGGGGATACGGCGACGAAGCCGGC  
GAACTTGCCCGCAACAGAAAACACGGGTTACGCTTGGCAGCTCCTACCCAACGGTATGAAGCCCGAATACCGCCCG  
TAACTCGAG

20 1 MASPDVKSAD TSKPAAPVV AEKETEVEKED APQAGSQGG APSTQGSQDM  
51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNQPADSS  
101 DSAPASNAP ANGGSNFRV DLANGVLIDG PSQNTITLTHC KGDSCNGDNL  
151 LDEEAPSKSE FENLNESERI EKYKDKGSD KFTNLVATAV QANGTNKYVI  
201 IYKDKSASS SARFRSARS RRSIPAEMPL IPVNQADTLI VDGEAVSLTG  
25 251 HSGNIFAPEG NYRYLTYGAR KLPGGSYALR VQGEPAKGEM LAGTAVYNGE  
301 VLHFHTENGR PYPTRGRFAA KVDGFSKSD GIIDSGDDLH MGTQKPKAAI  
351 DNGNGFKGTW ENGCGDVSGR FYGPAGEEVA GKYSYRPTDA EKGFGVVFAG  
401 KKEQDGS GCGSKSIQTF PQPDTSVING PDRFVGIPDP AGTTVCGGGA  
451 VYTVPHLSL PHWAAQDFAK SLQSFRLGCA NLKNRQGWQD VCAQAPQTFV  
30 501 HSFQAKOFFE RYFTPWQVAG NGSLAGTVTG YYEPVLKGDD RRTAQARFPI  
551 YGIPDDFISV PLPAGLRSGK ALVRIRQTGK NSGTIDNTGG THTADLSRFP  
601 ITARTTAIKG RFEGSRFLPY HTRNQINGGA LDGKAPILGY AEDPVELFFM  
651 HIQSGRLKT PSGYIRIGY ADKNEHPYVS IGRYMADKY LKLGQTSMQG  
701 IKAYMRQNPQ RLAEVLQNP SYIFFRELAG SSNDGPVGAL GTPLMGEYAG  
35 751 AVDRHYITLG APLFVATAHP VTRKALNRLI MAQDTGSAIK GAVRVDFYWG  
801 YGDEAGELAG KQKTTGYVWQ LLPNGMKPEY RP\*

**AG287-953**

40 ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG  
GTAAAAGAGATGCGCCACAGGCAGGTTCTCAAGGACAGGCGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA  
GTTTCGGCAGAAATACAGGCAATGGCGGTGCGGCAACACGGAACCAACCAAAATGAAGCAGGATACCGCAAAAT  
GATATGCCGCAAAATTCGCGCAATCCGCAATCAAACAGGGAACAACCAACCGCCGATTCTTCAGATTCCGCCCCC  
45 GCGTCAAACCCGTCACCTGCGAATGGCGGTAGCAATTTTGAAGGGTTGATTGTTGCTAATGGCGTTTGTATTGATGGG  
CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAAATGGTGATAATTTATTGGATGAAGAAGCACCG  
TCAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT  
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACATAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA  
TCTTCATCTGCGGATTTCAGGCGTTCTGACGGTCTGAGGAGGTGCGTTCTTGCGGAGATGCCGCTAATCCCCGTCAT  
50 CAGGCGGATACGCTGATTGTGATGGGGAAGCGGTGACCTGACGGGCGATTCCGGCAATATCTTCGCGCCCCGAAGGG  
AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCCGCGGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA  
AAAGGCGAAATGCTTGTGTCACGCGCCGTGTACAACGCGGAAGTGTGCTGATTTTCATACGGAACCGCCGTCCGTAC  
CCGACTAGAGGCGAGTTTGGCGCAAAAGTCGATTTCGGCAGCAAAATCTGTGGACGGCATATCGACAGCGGCGATGAT  
TTGCATATGGGTACGCAAAATTCAAAGCCGCATCGATGAAACCGGCTTAAGGGGACTTGGACGGAATGGCGGC  
55 GGGGATGTTTCCGGAAGGTTTACGGCCCGCGCGGAGGAGTGGCGGGAATAACGCTATCGCCCGACAGATGCG  
GAAAAGGGCGGATTTCGCGTGTTCGCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGAGCCACTACAAAGTG  
GACGAATATCACGCCAACGCCCGTTTCGCCATCGACCATTTCAACACCAGCACCAACGTCGGCGGTTTTTACGGTCTG  
ACCGGTTCCGTGAGTTCGACCAAGCAAAACGCGACGTTAAATCGACATCACCATCCCCGTGCAACCTGCAAGC  
GGTTCGCAACACTTTACCGACCACTGAAATCAGCCGACATCTTCGATGCCGCCCAATATCCGGACATCCGCTTTGTT  
60 TCCACCAAAATCAACTTCAACGGCAAAAACTGGTTTCCGTTGACGGCAACCTGACCATGCACGGCAAAACCGCCCC  
GTCAAACCTCAAAGCCGAAAAATCAACTGCTACCAAGCCGATGCGGCAAAACCGAAGTTTTCGCGCGGCGACTTCAG  
ACCACCATCGACCGACCAATGGGGCGTGACTACCTCGTTAAGTTGGTATGACCAAAAGCGTCCGCATCGACATC  
CAAATCGAGGCAGCCAAACAATAACTCGAG

65 1 MASPDVKSAD TSKPAAPVV AEKETEVEKED APQAGSQGG APSTQGSQDM  
51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNQPADSS  
101 DSAPASNAP ANGGSNFRV DLANGVLIDG PSQNTITLTHC KGDSCNGDNL  
151 LDEEAPSKSE FENLNESERI EKYKDKGSD KFTNLVATAV QANGTNKYVI

-9-

5

201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG  
 251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKGEM LAGTAVYNGE  
 301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKPKAAI  
 351 DGNFGKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGFGVVFAG  
 401 KKEQDGS GGGG GATYKVDEYH ANARFAIDHF NTSTNVGGFY GLTGSVEFDQ  
 451 AKRDGKIDIT IPVANLQSGS QHFTDHLKSA DIFDAAQYPD IRFVSTKFPNF  
 501 NGKKLVSVDG NLTMHGKTAP VKLKAKEFNC YQSPMAKTEV CGGDFSTTID  
 551 RTKWGVLDLV NVGMTKSVRI DIQIEAAKQ\*

10

ΔG287-961

15

20

25

30

35

40

ATGGCTAGCCCCGATGTTAAATCGGCGGACACGCTGTCAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG  
 GTAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCACACAAGGCAAGATATGGCGGCA  
 GTTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT  
 GATATGCCGCAAAATTCGCGCGAATCCGCAATCAAACAGGGAACAACCAACCGCCGATTCTTCAGATTCCGCCCCC  
 GCGTCAAACCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAAGGGTTGATTGGCTAATGGCGTTTTGATTGATGGG  
 CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGGATTCTTGTATGGTGATAATTTATTGGATGAAGAAGCACC  
 TCAAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT  
 ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACATAACAAATATGTCAATTTATAAAGACAAGTCCCGCTTCA  
 TCTTCATCTGCGCGATTTCAGGCGTTCTGACGCGTGCAGGAGGTGCGTTCTGCGGAGATGCCGCTAATCCCCGCTCAAT  
 CAGGCGGATACGCTGATTGTGATGGGGAAGCGGTGACGCTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGG  
 AATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCCGCGCGGATCGTATGCCCTCCGTTGTGCAAGGCGAACCGGCA  
 AAAGGCGAAATGCTTGCTGGCAGCGCCGTGTACAACGGCGAAGTGCTGCATTTTCATACGGAACCGCCGTCGCTAC  
 CCGACTAGAGGCAAGTTTGCCGCAAAAGTCGATTTCCGCGAGCAAAATCTGTGGACGGCATTATCGACAGCGCGGATGAT  
 TTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGATGGAAACGCGCTTAAAGGGGACTTGGACGGAATAATGGCGGC  
 GGGGATGTTTTCCGGAAGGTTTTACGGCCCGCGCGGAGGAGGTGGCGGAAAAATACAGCTATCGCCCGACAGATGCG  
 GAAAAGGGCGGATTTCGCGGTGTTTGCCGGCAAAAAGAGCAGGATGGATCCGGAGGAGGAGGACCAACGACGAC  
 GATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCATACAACAATGGCCAAAGAAATCAACGGTTTTCAAAGCTGGA  
 GAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGAC  
 TTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAAACAAACGTCGATGCC  
 AAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACAAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACT  
 GATGCCGCTCTGGATGCAACCACCAACGCTTGAATAAAATGGGAGAAAAATATAACGACATTTGCTGAAGAGACTAAG  
 ACAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGAT  
 ATCGCCGATTTCATTGGATGAAACCAACACTAAGGCAGACGAGCCGTCAAACCGCCAATGAAGCCAAACAGACGGCC  
 GAAGAAACCAACAAACGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAAGCCGAAGCTCCGCTGGC  
 ACAGCTAATACTGCAGCCGACAAGGCCGAGCTGCTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAAC  
 AAAGATAATATTGCTAAAAAAGCAACAGTGCAGCAGTGTACACCAGAGAAGAGTCTGACAGCAAAATTTGTGAGAATT  
 GATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAGGATACT  
 CGCCTGAACGGTTTTGGATAAAACAGTGTGACAGCTGCGCAAGAAACCGCCAAGGCTTGCAGAAACAGCCGCGCTC  
 TCCGGTCTGTTCCAACCTTACAACGTGGGTGCGTTCAATGTAACGGCTGCAGTCCGGCGGTACAAATCCGAATCGGCA  
 GTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCAAAAGCAGGCGTGGCAGTCCGCACTTCGTCGGGT  
 TCTTCCGACGCTACCATGTGCGCGTCAATTACGAGTGGTAACTCGAG

45

50

55

60

1 MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQ APSTQGSQDM  
 51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNQPADSS  
 101 DSAPASNPA ANGGSNFGVR DLANGVLIDG PSQNTLTHC KGDSCNGDNL  
 151 LDEEAPSKSE FENLNESERI EKYKDKGSD KFTNLVATAV QANGTNKYVI  
 201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG  
 251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKGEM LAGTAVYNGE  
 301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKPKAAI  
 351 DGNFGKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGFGVVFAG  
 401 KKEQDGS GGGG GATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE  
 451 DGTITKKDAT AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDKVKAAE  
 501 SEIEKLTTKL ADTDAALADT DAALDATNA LNKLGENTIT FAETKTNIIV  
 551 KIDEKLEAVA DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE  
 601 TKQNVDAKVK AAETAAGKAE AAAGTANTAA DKABAVAARKV TDIKADIATN  
 651 KDNIAKKANS ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH  
 701 DTRLNGLDKT VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK  
 751 SESAVAIGTG FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEW\*

	ELISA	Bactericidal
$\Delta$ G287-953-His	3834	65536
$\Delta$ G287-961-His	108627	65536

The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens (using 287-GST) for 919 and ORF46.1:

	Mixture with 287	Hybrid with $\Delta$ G287
919	32000	128000
ORF46.1	128	16000

Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained:

	919		ORF46.1	
Strain	Mixture	Hybrid	Mixture	Hybrid
NGH38	1024	32000	-	16384
MC58	512	8192	-	512
BZ232	512	512	-	-
MenA (F6124)	512	32000	-	8192
MenC (C11)	>2048	>2048	-	-
MenC (BZ133)	>4096	64000	-	8192

The hybrid proteins with  $\Delta$ G287 at the N-terminus are therefore immunologically superior to simple mixtures, with  $\Delta$ G287-ORF46.1 being particularly effective, even against heterologous strains.  $\Delta$ G287-ORF46.1K may be expressed in pET-24b.

The same hybrid proteins were made using New Zealand strain 394/98 rather than 2996:

10

 **$\Delta$ G287N2-919**

15

20

25

ATGGCTAGCCCCGATGTCAAGTCGGCGGACACGCTGTCAAAACCTGCCGCCCTGTTGTTTCTGAAAAAGAGACAGAG  
 GCAAAGGAAGATGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG  
 GTTTCGGAAGAAATACAGGCAATGGCGGTGCGGCAGCAACGACAAACCCAAAAATGAAGACAGGGGGCGCAAAAT  
 GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAT  
 ATGGAACCAAGCACCGGATGCCGGGAATCGGAGCAGCCGGCAACCAACCGGATATGGCAAATACGGCGGACGGA  
 ATGCAGGGTGACGATCCGTCGGCAGCGGGGAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAC  
 AATCAAACCGCGGTTCTCAAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTTTGAAGG  
 ACGAACGTGGGCAATCTGTTGTGATTGACGGGCCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT  
 AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAATCAGAATTTGAAAAATTAGTGATGCAGACAAATAAGT  
 AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG  
 GGAATCAATCAATATATTATCTTTTATAAACCTAAACCACTTCATTGCGCGATTAGGCGTTCTGCACGGTCGAGG  
 CGGTGCTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGC  
 CTGACGGGGCATTCCGGCAATATCTTCGCGCCCGAAGGGAATTACCGGTATCTGACTTACGGGGCGGAAAAATTGCC  
 GGCGGATCGTATGCCCTCCGTGTTCAAGGCGAACCTTCAAAAGGCGAAATGCTCGCGGGCACGGCAGTGTAACGGC  
 GAAGTGTGCAATTTTCATACGGAAACGGCCGTCCGTCCCGTCCAGAGGCAGGTTTGCCGCAAAAGTCGATTTCCGC  
 AGCAATCTGTGGACGGCATTATCGACAGCGCGATGGTTTCATATGGGTACGCAAAATTCAAAGCCGCCATCGAT  
 GGAAACGGCTTTAAGGGGACTTGGACGGAAATGGCGGGGGATGTTTCCGGAAAGTTTACGGCCCGGCCGCGAG

5 GAAGTGGCGGAAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTGCGCGTGTGTTGCCGGCAAAAAAGAG  
 CAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGCATCCAAACCTTTCCGCAACCCGACACATCCGTCATCAAC  
 GGCCCGGACCGGCCGGTCGGCATCCCCGACCCCGCCGGAACGACGGTCGGCGGCGGCGGGGCGCTCTATACCGTTGTA  
 CCGCACCTGTCCCTGCCCACTGGGCGGCGCAGGATTTCCGCAAAAGCCTGCAATCCTTCCGCTCGGCTGCGCCAAT  
 10 TTGAAAAACCGCCAAAGGCTGGCAGGATGTGTGCGCCCAAGCCTTTCAAACCCCGTCCATTCTTTTACGGCAAAACAG  
 TTTTGTGAACGCTATTTACGCCGTGGCAGGTTGCAGGCAACGGAAGCCTTGCCCGGTACGGTTACCGGCTATTACGAG  
 CCGGTGCTGAAGGGCGACGACAGGCGGACGGCACAAGCCCGCTTCCCGATTACGGTATTCCCGACGATTTATCTCC  
 GTCCCCCTGCCTGCCGGTTTGCGGAGCGGAAAAGCCCTTGTCGCGCATCAGGCAGACGGGAAAAAACAGCGGCACAATC  
 15 TTTGAAGGAAGCCGCTTCTCCCTACCACACGCGCAACCAAATCAACGGCGGCGCGCTTGACGGCAAAGCCCCGATA  
 CTCGGTTACGCCGAAGACCCCGTCGAATTTTTTTTATGCACATCCAAGGCTCGGGCCGCTCTGAAAACCCCGTCCGGC  
 AAATACATCCGCATCGCTATGCCGACAAAAACGAACATCCCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGC  
 TACCTCAAGCTCGGGCAGACCTCGATGCAGGCGCATCAAAGCCTATATGCGGCAAAATCCGCAACGCTCGCCGAAGTT  
 TTGGGTCAAACCCAGCTATATCTTTTCCGCGAGCTTGCCGGAAGCAGCAATGACGGTCCCGTCCGGCGCACTGGGC  
 20 ACGCCGTTGATGGGGGAATATGCCGCGCAGTCGACCGGCACTACATTACCTTGGGCGCGCCCTTATTGTGCGCCACC  
 CCCCATCCGGTTACCCGCAAAGCCCTCAACCGCCTGATTATGGCGCAGGATACCGGCAGCGCGATTAAAGGCGCGGTG  
 CGCGTGGATTATTTTGGGGATACGGCGACGAAGCCGGCGAACTTGCCGGCAACAGAAAACACGGGTACGTCTGG  
 CAGCTCTACCCAACGGTATGAAGCCGAATACCGCCCGTAAAAGCTT

20 1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQQG APSAQGGQDM  
 51 AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM  
 101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG  
 151 TNQAENNQTA GSQNPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNIL  
 201 THCKGDS CSG NNFLDEVLQ KSEFEKLSDA DKISNYKKDG KNDGKNDKFV  
 25 251 GLVADSVQMK GINQYIIFYK PKPTSFAFR RSARSRRSLP AEMPLIPVNV  
 301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAELKPGG SYALRVQGEF  
 351 SKGEMLAGTA VYNGEVLHFF TENGRPSR GRFAAKVDFG SKSVDGIIDS  
 401 GDGLHMGTOK FKAAIDGNF KGTWTENGSG DVSCKFYGPA GEEVAGKYSY  
 451 RPTDAEKGGF GVFAGKKEQD GSGGGGQSK SIQTFPPQDT SVINGPDRPV  
 30 501 GIPDPAGTTV GGGGAVYTVV PHLSLPHWAA QDFAKSLQSF RLGCANLKNR  
 551 QGWQDVCAQA PQTFVHSFQA KQFFERYFTP WQVAGNGSLA GTVTGYEYEV  
 601 LKGD RRRTAQ ARFPYIGIPD DFISVPLPAG LRSKALVRI RQTKNSGTI  
 651 DNTGGTHTAD LSRFPITART TAIKGRFEGS RFLPYHTRNQ INGGALDGKA  
 701 PILGYAEDPV ELFFMHQGS GRLKTPSGKY IRIGYADKNE HPYVSI GRYM  
 35 751 ADKGYLKLQ TSMQGIKAYM RQNPQLAEV LGQNPYSYIFF RELAGSSNDG  
 801 PVGALGTPLM GEYAGAVDRH YITLGAPLFV ATAHFVTRKA LNRLIMAQDT  
 851 GSAIKGAVRV DYFWGYGDEA GELAGKQKTT GYVWQLLPNG MKPEYRP\*

40 AG287NZ-953  
 ATGGCTAGCCCCGATGTCAAGTCGGCGGACACGCTGTCAAACCTGCCGCCCTGTGTTTCTGAAAAAGAGACAGAG  
 GCAAAGGAAGATGCCGCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG  
 GTTTCGGAAGAAAAATACAGGCAATGGCGGTGCGGCAGCAACGACAAACCCAAAAATGAAGACGAGGGGCGCAAAAT  
 45 GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAT  
 ATGGAAAACCAAGCACCGGATGCCGGGGAATCGGAGCAGCCGGCAAACCAACCGGATATGGCAAATACGGCGGACGGA  
 ATGCAGGGTGACGATCCGTGCGCAGGCGGGGAAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAAC  
 AATCAAACCGCCGTTCTCAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTGTTGGAAGG  
 ACGAAGCTGGGCAATTCTGTTGTGATTGACGGGCGCTCGCAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT  
 50 AGTGGCAATAATTTCTTGGATGAAGAAGTACAGTAAATCAGAAATTGAAAAATTAAGTGTGTCAGACAAAAAAGT  
 AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG  
 GGAATCAATCAATATATTATCTTTTATAAACCTAAACCCACTTCATTTGCGCGATTAGGCGTTCTGCACGGTCGAGG  
 CGGTGCTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTGATGGGGAAGCGGTCAGC  
 CTGACGGGGCATTCGGCAATATCTTCCGCCCGAAGGGAATTACCGGTATCTGACTTACGGGGCGGAAAAATGCCC  
 55 GCGGATCGTATGCCCTCCGTGTTCAAGGCGAACCTTCAAAGGCGCAATGCTCGCGGCGCAGGAGTGTACAACGGC  
 GAAGTGCTGCATTTTCATACGGAACACGGCGTCCGTCGCCGTCAGAGGCGAGTTTGCCGCAAAAGTCGATTTCGGC  
 AGCAAACTGTGTGGACGCGATTATCGACAGCGCGATGGTTTGTCATATGGGTACGCAAAATTCAAAGCCGCCATCGAT  
 GGAACCGCTTTAAGGGGACTTGACGGAATAATGGCGGCGGGGATGTTCCGGAAAGTTTACGGCCCGGCGCGGAG  
 GAAGTGGCGGAAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTTCGGCGTGTTCGGCGCAAAAAGAG  
 60 CAGGATGGATCCGGAGGAGGAGGCCACCTACAAAGTGGACGAATATCACGCCAACGCCCGTTTCGCCATCGACCAT  
 TTCAACACCAGCACCAACGTCCGCGGTTTTTACGGCTCTGACCGGTTCCGTCGAGTTTCGACCAAGCAAAACGCGACGGT  
 AAAATCGACATCACCATCCCCGTTGCCAACCTGCAAGCGGTTTCGCAACACTTTACCGACACCTGAAATCAGCCGAC  
 ATCTTCGATGCGCCCAATATCCGGACATCCGCTTGTTCACCAAAATTCAACTTCAACGGCAAAAAACTGTTTCC  
 65 GTTGACGGCAACCTGACCATGCACGGCAAAACCGCCCCCGTCAAACCTCAAAGCCGAAAAATTCAACTGCTACCAAAGC  
 CCGATGGCGAAAAACCGAAGTTTGGCGGCGGACTTCAGCACCACCATCGACCGCACCAATGGGGCGTGGACTACCTC  
 GTTAACGTTGGTATGACCAAAAGCGTCCGCATCGACATCCAAATCGAGGCGCAACCAATAAAAGCTT

51	AAVSEENTGN	GGAAATDKPK	NEDEGAQNDM	PQNAADTDSL	TPNHPTASNM
101	PAGNMENQAP	DAGESEQPAN	QPDMAANTADG	MQGDDPSAGG	ENAGNTAAQG
151	TNQAENNQTA	GSQNPASSTN	PSATNSGGDF	GRTNVGNVSV	IDGPSQNTIL
201	THCKGDSCSG	NNFLDEEVQL	KSEFEKLSDA	DKISNYKKDG	KNDGKNDKFV
251	GLVADSVQMK	GINQYIIFYK	PKPTSFAFR	RSARSRRSLP	AEMPLIPVNO
301	ADTLIVDGEA	VSLTGHSGNI	FAPEGNYRYL	TYGAEKLPBG	SYALRVQGEF
351	SKGEMLAGTA	VYNGEVLHFF	TENGRPSPSR	GRFAAKVDFG	SKSVDGIIDS
401	GDGLHMGTOK	FKAIDGNPF	KGTWTENGGO	DVSGKFYGPA	GEEVAGKYSY
451	RPTDAEKGGF	GVFAGKKEQD	GSGGGGATYK	VDEYHANARF	AIDHFNSTSN
501	VGGFYGLTGS	VEFDQAKRDG	KIDITIPVAN	LQSGSQHFTD	HLKSADIFDA
551	AQYPDIRFVS	TKFNFNGKKL	VSVDGNLTMH	GKTAPVKLKA	EKFNCYQSPM
601	AKTEVCCGDF	STTIDRTKWG	VDYLVNVGMT	KSVRIDIQIE	AAKQ*
<b>AG287NZ-961</b>					
15	ATGGCTAGCCCGATGTCAAGTCGGCGGACACGCTGTCAAACCTGCCGCCCTGTTGTTCTGAAAAAGAGACAGAG				
	GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCGCACAAAGGCGGTCAAGATATGGCGGCG				
	GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAACCCAAAAATGAAGACGAGGGGGCGCAAAAT				
	GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGCTTCGAATATGCCCGCGCGGAAAT				
20	ATGGAACCAAGCACCAGGATGCCGGGGAATCGGAGCAGCCGGCAACCAACCGGATATGGCAAATACGGCGGACGGA				
	ATGCAGGGTGACGATCCGTCGGCAGCGGGGAAATGCCGGCAATACCGGTGCCCAAGGTACAAATCAAGCCGAAAAC				
	AATCAAACCCCGGTTCTCAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTGGAAGG				
	ACGAACGTGGCAATCTGTTGTGATTGACGGGCCGTGCGAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT				
	AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAATCAGAATTTGAAAAATTAAGTGATGACACAAAATAAGT				
25	GGAATCAATCAATATATATCTTTTATAAACCTAAACCCACTTCATTGCGCGATTAGGCGTTCTGCACGGTCGAGG				
	CGGTGCTTCCGCGCGAGATGCCGCTGATTCCCGTCAATCAGCGGATACGCTGATTGTGCGATGGGGAAGCGGTCAGC				
	CTGACGGGGCATTCCGGCAATATCTTCGCGCCGAGGGAATACCGGTATCTGACTTACGGGGCGGAAAAATTGCC				
	GGCGGATCGTATGCCCTCCGTGTTCAAGCGCAACTTCAAAGCGCAAAATGCTCGCGGCACGGCAGTGTAACGGC				
30	GAAGTGCTGCATTTTCATACGGAAAAACGGCCGTCCGTCCCGTCCAGAGGCAGGTTTGCCGCAAAAGTCGATTTCCGC				
	AGCAATCTGTGGACGGCATTATCGACAGCGCGGATGTTTGATATGGGTACGCAAAATTCAAAGCCGCTTCGAT				
	GGAACCGGTTTAAAGGGACTTGGACGGAATGGCGCGGGGATGTTTCCGGAAGTTTACGGCCCGCGCGGAG				
	GAAGTGCGGGGAAAAATACAGCTATCGCCCAACAGATCGGGAAGGGCGGATTCGGCGTGTGCGCGCAAAAGAG				
	CAGGATGGATCCGGAGGAGGAGGAGCCACAAACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCC				
35	TACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACC				
	AAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTAAAGGTCTGGGTCTGAAAAAAGCTGCTGACTAACCTG				
	ACCAAAACCGTCAATGAAAAACAAACGCTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAAACA				
	ACCAAGTTAGCAGACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCTTGAATAAA				
	TTGGGAGAAAAATATAACGACATTGCTGAAGAGACTAAGACAAATATCGTAAAAATGATGAAAAATAGAAGCCGTG				
40	GCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTTCATTGGATGAAACCAACTAAGGCAGAC				
	GAAGCCGTCAAACCGCCAATGAAGCCAAACAGACGGCGGAAGAAACCAAAACGCTCGATGCCAAAGTAAAGCT				
	GCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAAGCCGAGCTGCTG				
	GCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGTG				
	TACACCAGAGAAGAGTCTGACAGCAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGC				
45	TTGGCTTCTGCTGAAAAATCCATTGCCGATCAGCACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGC				
	AAAGAAACCCGCCAAGGCCCTGTCAGAACAAAGCCGCTCTCCGCTGCTGTTCAAACCTTACAACGTGGGTGCGTTCAAT				
	GTAACGGCTGAGTCGGCGCTACAAATCCGAATCGGCAGTCGCCATCGGTACCGCTTCCGCTTACCGAAAACTTT				
	GCCGCCAAAGCAGGCGTGGCAGTCGGCACTTCGTCCGTTCTTCCGCGAGCTACCATGTGCGGCTCAATTACGAGTGG				
	TAAAGCTT				
50	1	MASPDVKSAD	TLSPKPAAPV	SEKETEAKED	APQAGSQGG
	51	AAVSEENTGN	GGAAATDKPK	NEDEGAQNDM	PQNAADTDSL
	101	PAGNMENQAP	DAGESEQPAN	QPDMAANTADG	MQGDDPSAGG
	151	TNQAENNQTA	GSQNPASSTN	PSATNSGGDF	GRTNVGNVSV
55	201	THCKGDSCSG	NNFLDEEVQL	KSEFEKLSDA	DKISNYKKDG
	251	GLVADSVQMK	GINQYIIFYK	PKPTSFAFR	RSARSRRSLP
	301	ADTLIVDGEA	VSLTGHSGNI	FAPEGNYRYL	TYGAEKLPBG
	351	SKGEMLAGTA	VYNGEVLHFF	TENGRPSPSR	GRFAAKVDFG
	401	GDGLHMGTOK	FKAIDGNPF	KGTWTENGGO	DVSGKFYGPA
	451	RPTDAEKGGF	GVFAGKKEQD	GSGGGGATND	DDVKKAAATVA
60	501	INGFKAGETI	YDIDRGTIT	KKDATAADVE	ADDFKGLGLK
	551	NENKQNVDAK	VKAABSEIEK	LTTKLADTDA	ALADTDAALD
	601	ENITFAEET	KTNIIVKIDEK	LEAVADTVDK	HAEAFNDIAD
	651	EAVKTANEAK	QTAETTKQNV	DAKVKAETA	AGKAEAAAGT
	701	VAAKVTDIKA	DIATNKDNIA	KKANSADVYT	REESDSKFPV
65	751	LDTRLASAEK	SIADHDTRLN	GLDKTVSDLR	KETRQGLAEQ
	801	NVGRFNVTA	VGGYKSESAV	AIGTGFRFTE	NFAAKAGVAV
	851	HVGVNYEW*			GTSSGSSAAY

**Example 3 – hybrids of  $\Delta$ G983**

Protein 983 has the following sequence:

	983	$\Delta$ G983
5	1 MRTPTPTPTK TFKPTAMALA VATTLSACLG	GGGGGTSAPD FNAGGTGIGS
	51 NSRATTAKSA AVSYAGIKNE MCKDRSMLCA	GRDDVAVTDR DAKINAPPPN
	101 LHTGDFPNPN DAYKNLINLK PAIEAGYTGR	GVEVGIVDTG ESVGSISFPE
	151 LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG	KDIEASFDDE AVIETEAKPT
	201 DIRHVKEIGH IDLVSHIIGG RSVDRPAGG	IAPDATLHIM NTNDETKNEM
10	251 MVAAIRNAWV KLGERGVRIV NNSFGTTSRA	GTADLFQIAN SEEQYRQALL
	301 DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK	NMLFIFSTGN DAQAQPNITYA
	351 LLPFYEKDAQ KGIITVAGVD RSGEKFKREM	YGEPTGTEPLE YGSNHCGITA
	401 MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP	IVTGTAALLL QKYPWMSNDN
	451 LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK	AMNGPASFPF GDFTADTKGT
15	501 SDIAYSFRND ISGTGGLIKK GGSQQLQHGN	NTYTGKTIIE GGSVLVLYGNN
	551 KSDMRVETKG ALIYNGAASG GSLNSDGIVY	LADTDQSGAN ETVHIKGSQ
	601 LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS	ARGKGAGYLN STGRRVPFLS
	651 AAKIGQDYSF FTNIETDGGG LASLDSVEKT	AGSEGDLSY YVRRGNAART
	701 ASAAAHSAPA GLKHAVEQGG SNLENLMVEL	DAESSATPE TVETAAADRT
20	751 DMPGIRPYGA TFRAAAAVQH ANAADGVRI	NSLAATVYAD STAAHADMQG
	801 RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW	EQGGVEGKMR GSTQTVGIAA
	851 KTGENTTAAA TLGMGRSTWS ENSANAKTDS	ISLFAGIRHD AGDIGYLKGL
	901 FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ	LGALGGVNVP FAATGDLTVE
	951 GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG	TLVGLAGLKL SQPLSDKAVL
25	1001 FATAGVERDL NGRDYYVTGG FTGATAATGK	TGARNMPHTR LVAGLGADVE
	1051 FGNWNGLAR YSYAGSKQYG NHSGRVGVGY	RF*

$\Delta$ G983 thus has the following basic sequence:

		TSAPD FNAGGTGIGS
30	NSRATTAKSA AVSYAGIKNE MCKDRSMLCA	GRDDVAVTDR DAKINAPPPN
	LHTGDFPNPN DAYKNLINLK PAIEAGYTGR	GVEVGIVDTG ESVGSISFPE
	LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG	KDIEASFDDE AVIETEAKPT
	DIRHVKEIGH IDLVSHIIGG RSVDRPAGG	IAPDATLHIM NTNDETKNEM
	MVAAIRNAWV KLGERGVRIV NNSFGTTSRA	GTADLFQIAN SEEQYRQALL
35	DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK	NMLFIFSTGN DAQAQPNITYA
	LLPFYEKDAQ KGIITVAGVD RSGEKFKREM	YGEPTGTEPLE YGSNHCGITA
	MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP	IVTGTAALLL QKYPWMSNDN
	LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK	AMNGPASFPF GDFTADTKGT
	SDIAYSFRND ISGTGGLIKK GGSQQLQHGN	NTYTGKTIIE GGSVLVLYGNN
40	KSDMRVETKG ALIYNGAASG GSLNSDGIVY	LADTDQSGAN ETVHIKGSQ
	LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS	ARGKGAGYLN STGRRVPFLS
	AAKIGQDYSF FTNIETDGGG LASLDSVEKT	AGSEGDLSY YVRRGNAART
	ASAAAHSAPA GLKHAVEQGG SNLENLMVEL	DAESSATPE TVETAAADRT
	DMPGIRPYGA TFRAAAAVQH ANAADGVRI	NSLAATVYAD STAAHADMQG
45	RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW	EQGGVEGKMR GSTQTVGIAA
	KTGENTTAAA TLGMGRSTWS ENSANAKTDS	ISLFAGIRHD AGDIGYLKGL
	FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ	LGALGGVNVP FAATGDLTVE
	GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG	TLVGLAGLKL SQPLSDKAVL
50	FATAGVERDL NGRDYYVTGG FTGATAATGK	TGARNMPHTR LVAGLGADVE
	FGNWNGLAR YSYAGSKQYG NHSGRVGVGY	RF*

$\Delta$ G983 was expressed as a hybrid, with ORF46.1, 741, 961 or 961c at its C-terminus:

**AG983-ORF46.1**

55	ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
	GCGAGTACTTTACGCCGGTATCAAGAACGAAATGTGCAAGACAGAACATGCTCTGTGCCGGTCGGGATGACGTTGCG
	GTTACAGACAGGATGCCAAAATCAATGCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
	TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
	ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
	AACTATACGGCGTATATGCGGAAGGAAGCGCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAG
60	GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCAT

ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG  
 AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACCTGGCGTGC  
 ATCGTCAATAACAGTTTGTGAACAACATCGAGGCGAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG  
 TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTAC  
 5 GGCAACCTGTCTTACCATCCGTAAATAAAACATGCTTTTCATCTTTTCGACAGGCAATGCAGGCACAAGCTCAGCCC  
 AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCGCAGT  
 GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTCGCGAATT  
 ACTGCCATGTGGTGCCTGTCCGGCACCTATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTGCCGGA  
 10 ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC  
 CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCCGGCTGGGGACTGCTG  
 GATGCGGTAAAGGCCATGAACGGACCCGCTCTTTCGTTCCGGTCCGGCACTTTACCGCGGATACGAAAGGTACATCCGAT  
 ATTGCCTACTCCTTCCGTAACGACATTTTCAGGCACGGGCGGCTGATCAAAAAAGGCGGCAACCTGCAACTGCAC  
 GGCAACAACACCTATACGGGCAAAACCATTTATCGAAGGCGGTTTCGTTGGTGTGTACGGCAACAACAAATCGGATATG  
 15 CGCGTCGAAACCAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT  
 CTGGCAGATACCGACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACG  
 CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCCGGCGCAAGCTGTACATGTCCGGCACGC  
 GGCAAGGGGCGAGGCTATCTCAACAGTACCGGACGAGCTTCCCTTCCAGTGCAGCGCAAAATCGGGCAGGATTAT  
 TCTTTCTTCAAAACATCGAAACCGGCGGCTGCTGGCTTCCCTCGACAGCGTCAAAAAACAGCGGCGAGTGAA  
 20 GCGACACGCTGTCTTATTATGTCGCTCGCGGCAATGCGGCACGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC  
 GGTCTGAAACACGCGTAGAACAGGCGGCGAGCAATCTGGAACCTGATGGTCAAGTGGATGCCTCCGAATCATCC  
 GCAACACCCGAGACGGTTGAAACTCGCGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTC  
 CGCGCAGCGGCGAGCCGTACAGCATGCGAATGCCGCGACGGTGTACGCATCTTCAACAGTCTCGCGGCTACCGTCTAT  
 CCGGCAATACCGCGCCCATGCGGATATGCGAGGACGCGCGCTGAAAGCGGTATCGGACGGTTGGACCAACCGGC  
 25 ACGGGTCTGCGCGTCTACGCGCAACCCCAACAGGACGGTGGAAACGTCGGGAAACAGGGCGGTTGAAGGCAAAATGCGC  
 GGCAGTACCCAAACCGTCCGCTTTCGCGCAAAACCGCGCAAAATACGACAGCAGCGGCCACACTGGGCATGGGACGC  
 AGCACATGGAGCGAAACAGTGCAAATGCAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGATGCGGGC  
 GATATCGGCTATCTCAAAGGCTGTCTCTTACGACGCTACAAAAACAGCATCAGCCCGCAGCACCGGTGCGGACGAA  
 CATGCGGAAGGCAGCGTCAACGGCAGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTCCGCAACG  
 30 GGATTTGACGGTTCGAAGGCGGTCTGCGCTACGACCTGCTCAAAACAGGATGCATTCCGCCGAAAGGCAAGTGTCTT  
 GGCTGGAGCGGCAACAGCCTCACTGAAGGACGCTGGTCCGATCTCGCGGTCTGAAGCTGTTCGCAACCTTTGAGCGAT  
 AAAGCCGCTCTGTTTGAACGGCGGGCGTGGAAACGCGACCTGAACGGACGCGACTACAGGTAACGGGCGGCTTTACC  
 GCGCGACTGCAGCAACCGGCAAGACGGGGGACGCAATATGCCGACACCCGCTCTGTTGCGCGCTGGGCGCGGAT  
 GTCGAATTCGGCAACGCGTGAACGCGTTGGCAGCTTACAGCTACGCGGTTCCAAACAGTACGGCAACCAACAGCGGA  
 35 CGAGTCGGCGTAGGCTACCGGTTCCGACGGTGGCGGAGGCACTGGATCCTCAGATTGGGCAACGATTCTTTTATC  
 CCGAGGTTCTCGACCGTCAGCATTTGCAACCGGACGGGAAATACCCTATTTCGGCAGCAGGGGGAACTTCCCGAG  
 CGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAACAGGCGGCCATTAAA  
 GGAATATCGGCTACATTGTCCGCTTTTCCGATCAGGGCACGAAGTCCATTCCCCCTTCGACAACCATGCCTCACAT  
 TCCGATTCTGATGAAGCGGTAGTCCCGTTGACGGATTTAGCTTTTACCGCATCCATTGGGACGGATACGAACACCAT  
 40 CCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGCGAGGGATATATACAGCTACGAC  
 ATAAAAGGCGTTGCCCAAAATATCCGCTCAACCTGACCGACAACCGCAGCAGCCGCAACCGCTTGGCGACCGTTTC  
 CACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGGCCACCCGATACAGCCCGAGCTGGAC  
 AGATCGGGCAATGCCGCGGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGCGCGGCAAGAGAAAT  
 45 GTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTATGCACGGCTTGGGTCTGCTTTCC  
 ACCGAAACAAGATGGCGCGCATCAACGATTGTCAGATATGGCGCAACTCAAAGACTATGCCGACGAGCCATCCGC  
 GATTGGGCGAGTCCAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCCATCCCCATC  
 AAGGGGATTGGAGCTGTTCCGGGAAAAATACCGCTTGGCGGCATCAGGCACATCCTATCAAGCGGTGCGAGATGGGC  
 50 CGGATCGATTGCCGAAAGGAAATCCGCGTCAAGGACAAATTTGCGGATCGCGGCATACGCCCAATACCCGTCGCCCT  
 TACCATTCCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAGAAAAACATCACTCTCAACCGTCCCGCG  
 TCAAACGGCAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGAAGGTAAGGGTTTCCG  
 AATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACTGA

1 MTSAPDFNAG GTGIGSNSRA TTAksAAVSy AGIKNEMCKD RSMLCAGRDD  
 51 VAVTDRDAKI NAPPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV  
 101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE  
 155 151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDLv SHIIGGRSVD GRPAGGIAPD  
 201 ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIvNNSF GTTSRAGTAD  
 251 LFQIANSEEQ YRQALLDYSg GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF  
 301 IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG  
 60 401 TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG  
 451 PASFPFGDFT ADTKGTSDIA YSFRNDISGT GGLIKKGGsQ LQLHGNNTYT  
 501 GKTIIEGGS L VLYGNKSDM RVETKGALiY NGAASGGSLN SDGIVYLADT  
 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK  
 601 GAGYLNSTGR RVPFLSAAKI GDYSFFTNi ETDGGLLASL DSVEKTAGSE  
 65 651 GDTLSYVVR GNAARTASAA AHSAPAGLKH AVEQGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAaVQHANAa DGVRIFNSLA  
 751 ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG



-15-

801 VEGKMRGSTQ TVGLIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVPFAAT GDLTVEGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG  
 951 LAGLKLSQL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR  
 1001 NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLD  
 1051 GGGGTGSSDL ANDSFIRQVL DRQHFEFDGK YHLFGSRGEL AERSGHIGLG  
 1101 KIQSHQLGNL MIQQAIAKGN IGYIVRFS DH GHEVHSPFDN HASHSDSDEA  
 1151 GSPVDGFSLY RIHWDGYEHH PADGYDGPQG GGYPAKPGAR DIYSYDIKGV  
 1201 AQNIRLNLTD NRSTGQRLAD RFHNAGSMLT QGVGDGFKRA TRYSPELDRS  
 1251 GNAAEAFNGT ADIVKNIIGA AGEIVGAGDA VQGISSEGSNI AVMHGLGLLS  
 1301 TENKMARIND LADMAQLKDY AAAAIRDWAV QNPNAAQGIE AVSNIFMAAI  
 1351 PIKIGIAVRG KYGLGGITAH PIKRSQMGAI ALPKGKSAVS DNFADAAYAK  
 1401 YPSPYHSRNI RSNLEQRYGK ENITSSTVPP SNGKNVKLAD QRHPKTGVFPF  
 1451 DGKGFNPFKE HVKYDTLEHH HHHH\*

**AG983-741**

ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA  
 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAGACAGAGCATGCTCTGTGCCGGTCGGGATGACGTTGCCG  
 GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCGAATCTGCATACCGGAGACTTCCAAACCCAAATGACGCA  
 TACAAGAATTGTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC  
 ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAAGTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA  
 AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTGACGATGAG  
 GCCGTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCAT  
 ATCGTCAATAACAGTTTGGGAACAACATCGAGGCGAGGCACTGCCGACCTTTTCCAAATAGCCAATTCCGAGGAGCAG  
 TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGTATCCGCCTGATGCAACAGAGCGATTAC  
 GGCAACCTGTCTTACCACATCCGTAATAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC  
 AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCGCAGT  
 GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACC GGGTACAGAACCGCTTGAGTATGGCTTCAACCATTCGCGAATT  
 ACTGCCATGTGGTGCTGTCCGCAACCTATGAAGCAAGCGTCCGTTTTCACCCGTACAAACCCGATTCAAATTGCCGGA  
 ACATCCTTTTCCGCAACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC  
 CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCCGCTGGGGACTGCTG  
 GATGCGGGTAAGGCCATGAACGACCCGCGTCTTTCCGTTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT  
 ATTGCCCTACTCCTTCCGTAACGACATTTTACGGCACGGGCGGCTGATCAAAAAGCGCGCAGCAACTGCAACTGCAC  
 GGCAACAACACCTATACGGGCAAAACATTATCGAAGGCGGTTTCGCTGGTGTGTGACGGCAACAACCAATTCGGATG  
 CGCGTCGAAACCAAGGTGCGCTGATTTATAACGGGCGGCTCCGGCGGCAGCTGAACAGCGACGGCATTGTCTAT  
 CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACG  
 CTGTACACACGTTTGGGCAAACCTGTGAAAGTGGACGGTACGGCGATTATCGGGCGCAAGCTGTACATGTCCGCACGC  
 GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTTTCCCTTCTGAGTGCCGCCAAATCGGGCAGGATTAT  
 TCTTTCTTCAAAACATCGAAACCGACGGCGGCTGTGGCTTCCCTCGACAGCTCGAAAAACAGCGGGGAGTAA  
 GGTCTGAAACACGCGCTAGAACAGGGCGGCAATCTGGAACCTGATGGTGAAGTGGATGCCCTCCGAATCATCC  
 GCAACACCCGAGACGGTTGAAACGCGGCAGCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTC  
 CGCGCAGCGGACGGCTACAGCATGCGAATGCCGCCGACGTTGTACGATCTTCAACAGTCTCGCCGCTACCGTCTAT  
 GCCGACAGTACCGCCGCCATGCGGATATGCAGGACGCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC  
 ACGGGTCTGCGCGTATCGCGCAAAACCAAGGACGGTGGAAACGTGGAAACAGGGCGGTGTTGAAGGCAAAATGCGC  
 GGCAGTACCCAAACCGTCGGCATTTGCCGCAAAACCGCGGAAATACGACAGCAGCGCCACACTGGGCATGGGACGC  
 AGCAGATGGAGCGAAAAACAGTGAATGCAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGC  
 GATATCGGCTATCTCAAAGGCTGTCTCTTACGGACGCTACAAAAACAGCATCAGCCCGAGCACCGGTGCGGACGAA  
 CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCCGCAACG  
 GGAGATTTGACGGTCAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCCGCCGAAAAAGGCAGTGTCTTG  
 GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTGCAGCTCGCGGTTGAAAGCTGTGCAACCTTTGAGCGAT  
 AAAGCCGTCTGTTTGAACGGCGGCGTGGAAACCGGACCTGTAACGGACGCGACTACACGGTAACGGGCGGCTTTACC  
 GGCGGACTGCAGCAACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGTCTGGTTGCCGGCTGGGCGCGGAT  
 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACACAGCGGA  
 CGAGTCGGCTAGGCTACCGGTTCTTCGAGGGATCCGGAGGGGGTGGTGTGCGCCCGGACATCGGTGCGGGGCTTGCC  
 GATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTGAGAAAAAC  
 GAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGTGACAGCTCAATACGGGCAAAATGAAG  
 TTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCTTTTACAGCCGAGCAAATACAAGATTCCGAGCATTCCGGG  
 AAGATGGTTGCGAAACGCCAGTTTCAAGTTCGGCGACATAGCGGGCAACATACATCTTTTGACAAGCTTCCCGAAGGC  
 GGCAGGGCGACATATCGCGGGACGGCGTTCGGTTTCAGACGATGCCGGCGGAAACTGACCTACACCATAGATTTCGCC  
 GCCAAGCAGGGAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAATGTGCACCTGGCCGCCCGGATATCAAG  
 CCGGATGGAAACGCCATGCCGTATCAGCGGTTCCGTCCTTTACAACCAAGCCGAGAAAGGCAGTTACTCCCTCGGT



ATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAAACCGTAAACGGCATACGCCATATCGGCCTT  
GCCGCCAAGCAACTCGAGCACCACCACCACCACCCTGA

5	1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAAVS	AGIKNEMCKD	RSMLCAGRDD
	51	VAVTDRDAKI	NAPPPNLHTG	DFPNPNDAYK	NLINLKPAIE	AGYTGRGVEV
	101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDIE
	151	ASFDDEAVIE	TEAKPTDIRH	VKEIGHIDL	SHIIGGRSVD	GRPAGGIAPD
	201	ATLHIMNTND	ETKNEMMVAA	IRNAWVKLGE	RGVRIVNNSF	GTTSRAGTAD
10	251	LFQIANSEEQ	YRQALLDYS	GDKTDEGIRL	MQQSDYGNLS	YHIRNKNMLF
	301	IFSTGNDQA	QPNTYALLPF	YEKDAQGII	TVAGVDRSGE	KFKREMYGEP
	351	GTEPLEYGSN	HCGITAMWCL	SAPYRASVRF	TRTNPIQIAG	TSFSAPIVTG
	401	TAALLLQKYP	WMSNDNLRTT	LLTTAQDIGA	VGVDKFGWG	LLDAGKAMNG
	451	PASFPFGDFT	ADTKGTSDIA	YSFRNDISGT	GGLIKKGGSQ	LQLHGNNTYT
15	501	GKTIIEGGS	VLYGNKSDM	RVETKGALY	NGAASGGS	SDGIVYLADT
	551	DQSGANETVH	IKGSLQLDGK	GTLYTRLGKL	LKVDGTAIIG	GKLYMSARGK
	601	GAGYLNSTGR	RVPFLSAAKI	GQDYSFFTNI	ETDGGLLASL	DSVEKTAGSE
	651	GDTLSYYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
	701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANA	DGVRIFNLSA
20	751	ATVYADSTAA	HADMQRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
	801	VEGKMRGSTQ	TVGIAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDSISLF
	851	AGIRHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQLGAL
	901	GGVNVFFAAT	GDLTVEGGLR	YDLLKQDAFA	EKGSALGWSG	NSLTEGTLVG
	951	LAGLKLSQLP	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAATGKTGAR
25	1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLF
	1051	SGGGGGVAAD	IGAGLADALT	APLDHKDKGL	QSLTLDQSVR	KNEKLKLAAQ
	1101	GAEKTYNGND	SLNTGKLNKD	KVSRFDFIRQ	IEVDGQLITL	ESGEFQVYKQ
	1151	SHSALTAFQT	EQIQDSEHSG	KMVAKRQFRI	GDIAGEHTSF	DKLPEGGRAT
	1201	YRGTAFGSDD	AGGKLYTTID	FAAQQNGKI	EHLKSPELNV	DLAAADIKPD
30	1251	GKRHAIVSGS	VLYNQAEKGS	YSLGIFGGKA	QEVAGSAEVK	TVNGIRHIGL
	1301	AAKQLEHHHH	HH*			

**AG983-961**

35	ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
	GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCCGGATGACGTTGCG
	GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA
	TACAAGAAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
	ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACCGCTATAACGAAAATTACAAA
40	AACATATACGGCGTATATGCGGAAGGAGCGCCTGAAGCAGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
	GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCAT
	ATTATTGGCGGGCGTTCGTTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
	AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCGC
	ATCGTCAATAACAGTTTGTGAACAACATCGAGGCGAGGCACTGCCGACCTTTTCCAAATAGCCAATTCCGAGGAGCAG
45	TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAACAGACGAGGGTATCCGCCGTGATGCAACAGAGCGATTAC
	GGCAACCTGTCTTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCAAGCTCAGCCC
	AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCCGAGGCGTAGACCGCAGT
	GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACC GGGTACAGAACCCTTGAGTATGGCTCCAACCATTGCGGAATT
	ACTGCCATGTGGTGCTGTGCGCACCCCTATGAAGCAAGCGTCCGTTTCAACCGTACAAACCCGATTCAAATTGCCGGA
50	ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCGTGATGAGCAACGACAAC
	CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCCGCGTGACAGCAAGTTCCGGCTGGGGACTGCTG
	GATGCGGGTAAGGCATGAACGGACCCGCGTCTTTCCGTTCCGGCGACTTTACCGCGGATACGAAAGGTACATCCGAT
	ATTGCCTACTCCTTCCGTAACGACATTTTCAGGCACGGGCGGCTGATCAAAAAGGCGGCAGCAACTGCAACTGCAC
	GGCAACAACACCTATACGGGCAAAACCATTTATCGAAGGCGGTTTCGCTGGTGTGTGTACGGCAACAACAAATCGGATATG
55	CGCGTCGAAACCAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
	CTGGCAGATACCGACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACG
	CTGTACACGCTTTGGGCAAACCTGCTGAAAGTGACGGTACGGCGATTATCGGGCGCAAGCTGTACATGTGGCAGCG
	GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGAGCTGTTCCCTTCTGAGTGCCGCCAAAATCGGGCAGGATTAT
	TCCTTTCTTACAAAACATCGAAACCGACGGCGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAA
60	GGCGACACGCTGTCTATTATGTCCGTCGCGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCGCC
	GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAACCTGATGCTTCACTCGAATGCTTCCGAAATCAGC

-17-

CATGCCGAAGGCAGCGTCAACGGCAGCGTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCCGCAACG  
 GGAGATTTGACGGTGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTGCGCCGAAAAAGGCAGTGCTTTG  
 GGCTGGAGCGGCAACAGCCTCACTGAAGGCAGCGTGGTGGGACTCGCGGGTCTGAAGCTGTGCAACCCCTTGAGCGAT  
 5 AAAGCCGTCCTGTTTGAACGGCGGGCGTGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC  
 GGCGCGACTGCAGCAACCGGCAAGACGGGGGCACGCAATATGCCGACACCCGCTGTTGGTGGCGGCTGGGCGCGGAT  
 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGGTTCCAAACAGTACGGCAACCACAGCGGA  
 CGAGTCGGCGTAGGCTACCGGTTCTCGAGGGTGGCGGAGGCACTGGATCCGCCACAAACGACGACGATGTTAAAAAA  
 GCTGCCACTGTGGCCATTGCTGCTGCCTACACAATGGCCAAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTAC  
 10 GACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTG  
 GGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAAACAAACGTCGATGCCAAAGTAAAGCT  
 GCAGAACTGTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG  
 GATGCAACCACCAACGCCTTGAATAAATTTGGGAGAAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTA  
 AAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTCA  
 15 TTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAATGAAGCCAAACAGACGGCCGAAGAAACCAAA  
 CAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGGCACAGCTAATACT  
 GCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAGCTGATATCGCTACGAACAAAGATAATATT  
 GCTAAAAAAGCAACAGTGCAGCGCTGTACACAGAGAAGAGTCTGACAGCAAAATTTGTGAGAAATGATGGTCTGAAC  
 GCTACTACCGAAAAATTTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAGGATACTCGCCTGAACGGT  
 20 TTGGATAAAACAGTGTGACACCTGCGCAAAGAAACCGCCAAAGCCCTTGCAAGCAAGCCGCGCTCTCCGGCTGTTTC  
 CAACCTTACAACGTGGGTGCGTTCAATGTAACGCTGCACTGCGGCGGTACAAATCCGAATCGGCAGTCGCCATCGGT  
 ACCGGCTTCCGCTTTTACCGAAAACTTTGCCGCCAAAGCAGCGTGGCAGTCGGCACTTCGTCCGGTTCTTCCGCAGCC  
 TACCATGTGCGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCTGA

1 MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD  
 25 51 VAVTDRDAKI NAPPNNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV  
 101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE  
 151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDL V SHIIGRSVD GRPAGGIAPD  
 201 ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIVNN SF GTTSRAGTAD  
 251 LFQIANSEEQ YRQALLDYS GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF  
 30 301 IFSTGNDQA QPNTYALLPF YEKDAQKGI TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG  
 401 TAALLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG  
 451 PASFPFGDFT ADTKGTSDIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT  
 501 GKTIEGGS L VLYGNKSDM RVETRGALIIY NGAASGGS LN SDGIVYLADT  
 35 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK  
 601 GAGYLNSTGR RVPFLSAAKI GDYSFFTNI ETDGGLLASL DSVEKTAGSE  
 651 GDTLSEYVRR GNAARTASAA AHSAPAGLKH AVEQGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANA DGVRIFFNSIA  
 751 ATVYADSTAA HADMQRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG  
 40 801 VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGLTVG  
 951 LAGLKLSQL SDKAVLFATA GVERDLNDRD YTVTGFTGA TAATGKTGAR  
 1001 NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE  
 45 1051 GGGGTGSATN DDDVKKAATV AIAAAYNNGQ EINGFKAGET IYDIDEDGTI  
 1101 TKKDATAADV EADDFKGLGL KKVVTNLTKT VNEKNQNVDA KVKAASEIE  
 1151 KLTTKLADTD AALADTDAAL DATTNALNKL GENITTFEE TKTNIWKIDE  
 1201 KLEAVADTVD KHAEAFNDIA DSLDETNTKA DEAVKTANEA KQTAETKQN  
 1251 VDAKVAAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI  
 50 1301 AKKANSADV TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL  
 1351 NGLDKTVSDL RKETROGLAE QAALSGLFQP YNVGRFNVTA AVGGYKSESA  
 1401 VAIGTGFRFT ENFAAKAGVA VGTSSGSAA YHVG VNYEWL EHHHHHH\*

55 AG983-961c  
 ATGACTTCTGCGCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACACAGCGAAATCAGCA  
 GCAGTATCTTACGCCGGTATCAAGAACGAAATGTGCAAGACAGAAGCATGCTCTGTGCCGGTGGGATGACGTTGCG  
 GTTACAGACAGGGATGCCAAATCAATGCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCA  
 60 TACAAGAATTTGATCAACCTCAAACCTCAATGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC  
 ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA  
 AACTATACGGCTATATGCGGAAGGAAGCGCTGAAGACGGAGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAG  
 GCCGTATAGACTGAAGCAAGCCGACGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCAT  
 ATTATTGGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTCGCCCGGATGCGACGCTACACATAATGAATACG  
 AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGGCG  
 65 ATCGTCAATAACAGTTTTTGGAAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG  
 TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC  
 GGCAACCTGTCTTACCACATCCGTAATAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC

AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCGAGGCGTAGACCGCAGT  
 GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGCCGAATT  
 ACTGCCATGTGGTGCTGTGCGCACCCCTATGAAGCAAGCGTCCGTTTACCCGTACAAAACCCGATTCAAATTTGCCGGA  
 5 ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGTCAGAAAATACCCGTGGATGAGCAACGACAAC  
 CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCCGGCTGGGGACTGCTG  
 GATGCGGGTAAGGCCATGAACGGACCCGCTCCTTTCCGTTCCGGCGACTTTACCGCCGATACGAAAAGGTACATCCGAT  
 ATTGCCTACTCCTTCCGTAAACGACATTTAGGCACGGCGCGCTGATCAAAAAAGGGCGAGCAACTGCAACTGCAC  
 10 GGCAACAACACCTATACGGGCAAAACCATTTATCGAAGGCGGTTCCGTTGGTGTGTACGGCAACAACAAATCGGATATG  
 CGCGTCGAAACCAAAGGTGCGCTGATTATTAACGGGGCGGCATCCGGCGGCAGCTGAAACAGCGACGGCATTGTCTAT  
 CTGGCAGATACCGACCAATCCGGCGCAAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACG  
 CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGCGCGCAAGCTGTACATGTCGGCACGC  
 GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCTGAGTGCAGCGCAAAATCGGGCAGGATTAT  
 TCTTTCTTCAAAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCGAAAAAAGCGCGGAGTGA  
 15 GGCGACACGCTGTCTTATGTCCGTGCGGCAATGCGGCACGGACTGCTTCCGCGAGCGGCACATTCCGCGCCCGCC  
 GGTCTGAAACACGCGTAGAACAGGGCGGCAATGCGGAAACCTGATGGTTCGAACTGGATGCTCCGAATCATCC  
 GCAACACCCGAGACGGTTGAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCCACGCGCAACTTTTC  
 CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTAT  
 CCGGACAGTACCGCCCGCCATGCCGATATGCAGGGACGCCGCTGAAAGCGGTATCGGACGGGTTGGACCAACACGGC  
 20 ACGGGTCTGCGCGTCATCGCGCAAAACCAACAGGACGGTGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGC  
 GGCAGTACCCAAACCGTCGGCATTGCCCGCAAAACCGGCAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGC  
 AGCAGATGGAGCGAAACAGTGCAAATGCAAAACCGACAGCATAGTCTGTTTGCAGGCATACGGCAGCATGCGGGC  
 GATATCGGCTATCTCAAAGGCCTGTTCTCTTACGGACGCTACAAAAACAGCATCAGCCGACGACCCGGTGGCGAGCAA  
 CATGCGGAAGGCAGCGTCAACGGCAGCGTGTGAGCTGGGCGCACTGGCGGTTCAACGTTCCGTTTGGCGCAACG  
 25 GGAGATTTGACGGTGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCCGCCAAAAAGCGGATGCTTTG  
 GGTGGAGGCGCAACAGCCTCACTGAAGGCAGCTGGTGGGACTCGCGGTCTGAAGCTGTGCAACCCCTTGAGCGAT  
 AAAGCCGTCCTGTTTGAACGGCGGGCGTGAACCGGACCTGAACGGACGCGACTACAGGTAACGGGCGGCTTTACC  
 GCGCGACTGCAGCAACCGGCAAGACGGGGCGCAATATGCCGCACACCCGCTGTTGTCGGCGCTGGGCGCGGAT  
 GTCGAATTCGGCAACGGCTGGAACGGCTTGGCAGCGTTACAGCTACGCCGTTTCAAACAGTACGGCAACACAGCGGA  
 30 CGAGTCCGCGTAGGCTACCGGTTTCTCGAGGGTGGCGGAGGCACTGGATCCGCCCAAAACGACGAGATGTTAAAAAA  
 GCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACATCTAC  
 GACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTG  
 GGTCTGAAAAAAGTCTGTGACTAACCTGACCAAAACCGTCAATGAAACAAACAAACGTCGATGCCAAAGTAAAGCT  
 GCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG  
 35 GATGCAACCACCAACGCCCTTGAATAAATTGGGAGAAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTA  
 AAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCCGAAGCATTAACGATATCGCCGATTC  
 TTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCGCCAATGAAGCCAAACAGACGGCCGAAGAAACCAA  
 CAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGGCACAGCTAATACT  
 40 GCAGCCGACAAAGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATT  
 GCTAAAAAAGCAAACAGTGGCGACGTGTACACCGAGAGAGTCTGACAGCAAAATTTGTGAGAATTGATGGTCTGAAC  
 GCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAGGATCTCGCTGAACGGT  
 TTGGATAAAACAGTGTGAGACCTGCGCAAGAAACCCGCCAAGGCCTTGCAGAACAGCCGCGCTCTCCGGTCTGTTTC  
 CAACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCTGA

1 MTSAPDFNAG GTGIGSNSRA TTAksAAVSy AGIKNEMCKD RSMLCAGRDD  
 45 51 VAVTDRDAKI NAPPNNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV  
 101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE  
 151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDL V SHIIGGRSVD GRPAGGIAPD  
 201 ATLHIMTND ETKNEMMVA A IRNAWVKLGE RGVRIVNSF GTTSRAGTAD  
 251 LFQIANSEEQ YRQALLDYS GDKTDEGIRL MQQSDYGNLS YHIRKNMLF  
 50 301 IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG  
 401 TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVD SKFGWG LLDAGKAMNG  
 451 PASFFPGDFT ADTKGTS DIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT  
 55 501 GKTIIEGSL VLYGNKSDM RVETKGALIY NGAASGSLN SDGIVYLADT  
 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIG GKLYMSARGK  
 601 GAGYLNSTGR RVEFLSAKI GDYSFFTNI ETDGGLLASL DSVEKTAGSE  
 651 GDPLSYVRR GNAARTASAA AHSAPAGLKH AVEQGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANA DGVRIWFNSLA  
 751 ATVYADSTAA HADMQRRLL AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG  
 60 801 VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSY RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVFFAAT GDLTVEGLR YDLLKQDAFA EKGSLGWSG NSLTEGTLVG  
 951 LAGLKSQPL SDKAVLFATA GVERDLNGRD YTVTGGFTGA TAATGKTGAR  
 1001 NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLF  
 65 1051 GGGGTGSATN DDDVKKAATV AIAAAYNNGQ EINGFKAGET IYDIDEDGTI  
 1101 TTKDATAADV EADDFKGLGL KRVVTNLTKT VNEKNQNVDA KVKAASEIE  
 1151 KLTTKLADTD AALADTDAAL DATTNALNKL GENITTFABE TKTNIWKIDE

1201 KLEAVADTVD KHAEAFNDIA DSI DETNTKA DEAVKTANEA KQTAEETKQN  
 1251 VDAKVKAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI  
 1301 AKKANSADVY TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL  
 1351 NGLDKTVSDL RKETRQGLAE QAALSGLFQP YNVGLEHHHH HH\*

5

**Example 4 – hybrids of  $\Delta G741$** 

Protein 741 has the following sequence:

1 VNRTAFCCLS LTTALILTAC SSGGGGVAAD IGAGLADALT APLDHRDKGL  
 51 QSLTLDQSVR KNEKLKLAAG GAETTYGNGD SLNTGKLKND KVSRLFDFIRQ  
 101 IEVDGQLITL ESSEFQVYKQ SHSALTAFQT EQIQDSEHSG KMAKRFQRI  
 151 GDIAGEHTSF DKLPEGGGRAT YRGTAFGSDD AGGKLTYYTID FAAKQNGKI  
 201 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA  
 251 QEVAGSAEVK TVNGIRHIGL AAKQ\*

10

15

 $\Delta G741$  thus has the following basic sequence:

VAAD IGAGLADALT APLDHRDKGL  
 QSLTLDQSVR KNEKLKLAAG GAETTYGNGD SLNTGKLKND KVSRLFDFIRQ  
 IEVDGQLITL ESSEFQVYKQ SHSALTAFQT EQIQDSEHSG KMAKRFQRI  
 GDIAGEHTSF DKLPEGGGRAT YRGTAFGSDD AGGKLTYYTID FAAKQNGKI  
 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA  
 QEVAGSAEVK TVNGIRHIGL AAKQ\*

20

 $\Delta G741$  was fused directly in-frame upstream of proteins 961, 961c, 983 and ORF46.1:

25

 **$\Delta G741-961$** 

ATGGTTCGCCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCACCGCTCGACCATAAAGACAAAGGTTTGCAG  
 TCTTTGACGCTGGATCAGTCCGTGAGAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAAGTTATGGA  
 AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTGAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG  
 30 GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCGCGCTTAACCGCCTTTTCAG  
 ACCGAGCAAATACAAGATTTCGGAGCATTCGCGGAAGATGGTTGCGAAACGCCAGTTTCAAGATCGGGCAGATAGCGGGC  
 GACATACATCTTTTGACAAGCTTCCGGAAGCGCGGAGGGCGACATATCGCGGGACGGCGTTTCGGTTTCAGACGATGCC  
 GGCGGAAAACTGACCTACACCATAGATTTCGCGGCCAAGCAGGGAAACGGCAAAATCGAACATTTGAAATCGCCAGAA  
 CTCAATGTCGACCTGGCCGCCGCGGATATCAAGCCGATGGAACGCCATGCGGTCATCAGCGGTTCCGTCTCTTTAC  
 AACCAAGCCGAGAAAGGCAGTTACTCCTTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGCGGAAGTG  
 35 AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGTGGCGGAGGCACTGGATCCGCCACA  
 AACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCTTACAAACATGGCCAAAGAAATCAACGGTTTC  
 AAAGCTGGAGAGACCATCTACGACATTGATGAAGACCGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAA  
 GCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCTGACTAACCAGCAAAACCGTCAATGAAAAACAAACAAAC  
 40 GTCGATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTA  
 GCAGATACTGATGCCGCTCTGGATGCAACCAACGCTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAA  
 GAGACTAAGACAAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCA  
 TTCAACGATATCGCCGATTCAATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCCAATGAAGCCAAA  
 CAGACGGCCGAGAAACCAACAAACGTCGATGCCAAGTAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCT  
 45 GCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCTGATATC  
 GCTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGTGTACACAGAGAAGAGTCTGACAGCAAAATTT  
 GTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT  
 CACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACAGCTGCGCAAAGAACCCGCAAGGCCTTGCAAGCAA  
 GCCCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTTCGGTTCAATGTAACGGCTGCAAGTCGGCGGCTACAAATCC  
 50 GAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTCGGCAC  
 TCGTCCGGTTCTTCCGCGAGCTACCATGTCGGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCTGA

50

1 MVAADIGAGL ADALTAFLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAETK  
 51 YNGDSLNTG KLKNDKVSF DFIRQIEVDG QLITLESSEF QVYKQSHSAL  
 101 TAFQTEQIQD SEHSGKMAK RFIRIGDIAE EHTSFDKLPE GGRATYRGTA  
 151 FGSDDAGGKL TYTIDFAAQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
 201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL  
 251 EGGGTGSAT NDDDVKKAAT VAIAAAYNNG QEINGFKAGE TIYDIDEDGT  
 301 ITRKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNV AKVKAASEI  
 351 EKLTTKLADT DAALADTDA LDATTNALNK LGENITTFAE ETKTNIVKID  
 60 401 EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAEETKQ

55

60

-20-

451 NVDAKVKAAE TAAGKAEAAA GTANTAADKA EAVA AKVTDI KADIATNKDN  
501 IAKKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR  
551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGRFNV T AAVGGYKSES  
601 AVAIGTGFRF TENFAAKAGV AVGTSSGSSA AYHVG VNYEW LEHHHHHHH\*

5

**AG741-961c**

ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACC GCACCGCTCGACCATAAAGACAAAGGTTTCAG  
TCTTTGACGCTGGATCAGTCCGT CAGGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA  
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGT CAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG  
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG  
ACCGAGCAAATACAAGATTCCGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTT CAGAATCGCGGACATAGCGGGC  
GAACATACATCTTTGACAAGCTTCCCGAAGCGGCAGGGCGACATATCGCGGGACGGCGTTCCGTT CAGACGATGCC  
GGCGGAAAACTGACCTACACCATAGATTTCCGCCCAAGCAGGTAACCTGACCAAAACCGTCAATGAAAAACAAACAAAC  
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGATGGAAAAACGCCATGCCGT CATCAGCGTTCCGTCCTTTAC  
AACCAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAAGCCAGGAAGTTGCCGCGCAGCGCGGAAGTG  
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGTGCGCGAGGCACTGGATCCGCCACA  
AAGCAGCAGCATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCTTACAACAATGGCCAAGAAATCAACGGTTTC  
AAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAA  
GCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCTGTGACTAACCCTGACCAAAACCGTCAATGAAAAACAAACAAAC  
GTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTA  
GCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATGGGAGAAAATATAACGACATTTGCTGAA  
GAGACTTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGCTGATACCTCGACAAGCATGCCGAAGCA  
TTCAACGATATCGCCGATTCAATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCGCCAATGAAGCCAAA  
CAGACCGCCGAAGAAACCAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCT  
GCCGCTGGCACAGCTAATACTGCAGCCGACAAAGCCGAAGCTGTCTGCTGCAAAAGTTACCGACATCAAAGCTGATATC  
GCTACGAACAAAGATAATATTGCTAAAAAGCAAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTT  
GTCAGAAATTGATGCTGTAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT  
CACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAGAAACCCGCCAAGGCTTGCAGAACAA  
GCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCTGA

1 MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLA AQGA EKT  
51 YNGDSLNTG KLKNDKVS RF DFIRQIEVDG Q LITLES GEF QVYKQSHSAL  
101 TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA  
151 FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL  
251 EGGGTGSAT NDDDVKKAAT VAIAAAYNNG QEINGFKAGE TIYDIDEDGT  
301 ITKKDATAAD VEADDFKGLG LKKVVNTLTK TVNENKQNV D AKVKA ESEI  
351 EKLTTKLADT DAALADTDA A LDATTNALNK LGENITTFAE ETKTNIVKID  
401 EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAETKQ  
451 NVDAKVKAAE TAAGKAEAAA GTANTAADKA EAVA AKVTDI KADIATNKDN  
501 IAKKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR  
551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGL EHHH HHH\*

45

**AG741-983**

ATGGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACC GCACCGCTCGACCATAAAGACAAAGGTTTCAG  
TCTTTGACGCTGGATCAGTCCGT CAGGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA  
AACGGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGT CAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG  
GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTCAG  
ACCGAGCAAATACAAGATTCCGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTT CAGAATCGCGGACATAGCGGGC  
GAACATACATCTTTTGACAAGCTTCCCGAAGCGGCAGGGCGACATATCGCGGGACGGCGTTCCGTT CAGACGATGCC  
GGCGGAAAACTGACCTACACCATAGATTTCCGCCCAAGCAGGGAACCGGCAAAATCGAACATTTGAAATCGCCAGAA  
CTCAATGTCGACCTGGCCGCCGCCGATATCAAGCCGATGGAAAAACGCCATGCCGT CATCAGCGTTCCGTCCTTTAC  
AACCAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAAGCCAGGAAGTTGCCGCGCAGCGCGGAAGTG  
AAAACCGTAAACGGCATACGCCATATCGGCCTTGCCGCCAAGCAACTCGAGGGATCCGGCGGAGCGGCACTTCTGCG  
CCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTAC  
GCCGTATCAAGAACGAAATGTCAAAGACAGAAGCATGCTCTGTGCCGCTCGGGATGACGTTGCCGTTACAGACAGG  
GATGCCAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAAATTTG  
ATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCC  
GTCGCGCAGCATCTCTTTCCGAACTGTATGGCAGAAAAAGAACAGGCTATAACGAAAATTACAAAACTATACGGCG  
TATATGCGGAAGGAAGCGCCTGAAGACGGGCGGTAAAGACATTGAAGCTTCTTTCCAGCATGAGGCGCTTATAGAG  
ACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGG  
CGTTCCGTTGACGCGCAGACCTGCAGGCGGTATGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACC  
AAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGCGCAACGTGGCGTGCGCATCGTCAATAAC  
AGTTTGTGAACAAATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCG  
TTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCC

65

TACCACATCCGTAATAAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCC  
 CTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCGCAGTGAGAAAAAGTTC  
 AAACGGGAAATGTATGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTCGCGGAATTACTGCCATGTGG  
 5 TGCTGTGTCGGCACCCCTATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTGCCGGAACATCCTTTTCC  
 GCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACG  
 TTGCTGACGACGGCTCAGGACATCGGTGTCAGTCCGGCTGGACAGCAAGTTCGGCTGGGAGCTGCTGGATGCGGGTAAG  
 GCCATGAACGGACCCCGCTCTTTCCGTTCCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCTACTCC  
 TTCCGTAACGACATTTACGGCACGGCGGCCTGATCAAAAAGGCGGCAGCCAACCTGCAACTGCACGGCAACAACACC  
 10 TATACGGGCAAAACCATTTATCGAAGGCGGTTCGCTGGTGTGTACGGCAACAACAATCGGATATGCGCGTCGAAACC  
 AAAGGTGCGCTGATTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTTGTCTATCTGGCAGATACC  
 GACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACGCTGTACACACGT  
 TTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGGCGGCAAGCTGTACATGTCCGCGCAGCGCAAGGGGGCA  
 GGCTATCTCAACAGTACCGGACGAGCTGTTCCCTTCTGAGTGCCGCCAAAATCGGGCAGGATTATTTCTTTCTCACA  
 15 AACATCGAAACCGACGGCGGCCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCGACACGCTG  
 TCCTATTATGTCCGTCGCGGAATGCGGCACGAGTGTCTCGGCAGCGGCACATTCGCGCGCCCGCGGTCTGAAACAC  
 GCCGTAGAACAGGGCGGCAGCAATCTGGAACCTGATGGTCGAACTGGATGCCTCCGAAATCATCCGCAACACCCGAG  
 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTCCGCGCAGCGGCA  
 20 GCCGTACAGCATGCGAATGCCCGCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGCTCTATGCCGACAGTACC  
 GCCGCCATGCCGATATGCAGGGACGCCGCTGAAAGCCGTATCGGACGGGTGGACCAACCGGCACGGGCTCTGCGC  
 GTCATCGCGCAAAACCAACAGGACGGTGGAACTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAGTACCCAA  
 ACCGTTCGGCATTGCCGCGAAAACCGCGCAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCACATGGAGC  
 GAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTAT  
 25 CTCAAAGGCTGTTCTCCTACGGACGCTACAAAACAGCATCAGCCGACGACCGGTGCGGACGAAACATGCGGAAGGC  
 AGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTTCGCGCAACGGGAGATTGAGCG  
 GTCGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTTCGCCGAAAAAGGCAGTGCTTTGGGCTGGAGCGGC  
 AACAGCCTCACTGAAGGCACGCTGGTTCGGACTCGCGGGTCTGAAGCTGTGCAACCCCTTGAGCGATAAAGCCGTCCTG  
 TTTGCAACGGCGGGCGTGAACGCGACCTGAACGCGACGCTACACGGTAACGGGCGGCTTTACCGCGCGGACTGCA  
 30 GCAACCGGCAAGACGGGGGCACGCAATATGCCGCAACCCGCTCTGGTTGCCGGCCTGGGCGGGATGTCGAATTCGGC  
 AACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGTTTCCAAACAGTACGGCAACCACAGCGGACGAGTCCGCGTA  
 GGCTACCGGTTCTCTGAGCACCACCACCACCACCTGA

1 MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAETK  
 51 YNGNDSLNTG KLKNDKVSFR DFIRQIEVDG QLITLESGEF QVYKQSHSAL  
 101 TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLE GGRATYRGTA  
 151 VSGDDAGGKL TYTIDFAAQK GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
 201 FGSVSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL  
 251 ESGGGGTSAP PDFNAGGTGI GSNSRATTAK SAAVSYAGIK NEMCKDRSML  
 301 CAGRDDVAVT DRDAKINAPP PNLHTGDFPN PNDAYKNLIN LKPAIEAGYT  
 351 GRGVEVGIVD TGESVGSISF PELYGRKEHG YNENYKNYTA YMRKEAPEDG  
 401 GGKDIEASFD DEAVIETEA PTDIRHVKEI GHIDLVSII GGRSVDGRPA  
 451 GGIAPDATH IMNTNDETKN EMMVAAIRNA WVKLGERGVR IVNNSFGTTS  
 501 RAGTADLFQI ANSEEQYRQA LLDYSGGDKT DEGIRLMQQS DYGNLSYHIR  
 551 NKNMLFIFST GNDAQAPNT YALLPFYEKD AQKGIIIVAG VDRSGEKFPR  
 601 EMYGEPGTEP LEYGSNHCGI TAMWCLSAFY EASVRFTRTN PIQIAGTSFS  
 451 651 APIVTGTAAL LLQKYPWMSN DNLRTLLLT AQDIGAVGVD SKFGWGLLDA  
 701 GKAMNGPASF PFGDFTADTK GTSDIAYSFR NDISGTGLI KKGSQQLQLH  
 751 GNNTYTGTI IEGGSLVLYG NNKSDMRVET KGALIYNGAA SGGSLNSDGI  
 801 VYLADTDQSG ANETVHIKGS LQLDGKGLY TRLGKLLKVD GTAIIGKLY  
 851 MSARGKGAGY LNSTRRVVF LSAKIGQDY SFFTNIETDG GLLASLDSVE  
 501 901 KTAGSEGDTL SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV  
 951 ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR  
 1001 IFNSLAATVY ADSTAAHADM QGRRRLKAVSD GLDHNGTGLR VIAQTQDDGG  
 1051 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT  
 1101 DSISLFAGIR HDAGDIGYK GLFSYGRYKN SISRSTGADE HAEGSVNGTL  
 551 1151 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKG ALGWSGNSLT  
 1201 BGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT  
 1251 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV  
 1301 GYRFLHHHH HH\*

60

**AG741-ORF46.1**

ATGGTTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTCAG  
 TCTTTGACGCTGGATCAGTCCGTACAGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAAGTTATGGA  
 AACGGTGACAGCCTCAATACGGGCAAAATGAAGAACGACAAGGTACGCCGTTTCGACTTTATCCGCCAAATCGAAGTG  
 651 GACGGGACGCTCATTTACCTTGGAGAGTGGAGATTCCAAGTATACAAACAAGCCATTCCGCTTAACCGCTTTTCAG  
 ACCGAGCAAAATACAAGATTTCGAGCATTCGGGGAAGCATGGTTTCGAAACGCCAGTTTCAGATTCGGCGACATAGCGGC  
 GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCAGGGCGACATATCGCGGGACGGCTTCGGTTTCAGACGATGCC

5 GGC GGAAA ACTGACCTACACCATAGATTTGCGCCGCAAGCAGGGAACGGCAAAATCGAACATTTGAAATCGCCAGAA  
 CTCAATGTCGACCTGGCCGCCGCGGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTAC  
 AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGCGGAAGTG  
 AAAACCGTAAACGGCATAACGCATATCGGCCTTGCCGCCAAGCAACTCGACGGTGCGGGAGGCACTGGATCCTCAGAT  
 10 TTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTTGAACCCGACGGGAAATACCACCTATTCCGGC  
 AGCAGGGGGGAACTTGCCGAGCGCAGCGGCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATG  
 ATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCC  
 TTCGACAACCATGCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGAATTTAGCCTTTACCGCATCCAT  
 TGGGACGGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGCGGTATCCCGCTCCCAAAGGCGCG  
 15 AGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCAGGA  
 CAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACC  
 CGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATC  
 ATCGGCGCGCAGGAGAAATGTGCGGCGCAGCGGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATGTCTGTATG  
 CACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTGGCAGATATGGCGCAACTCAAAGAC  
 20 TATGCCGACGACGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCGACAAGGCATAGAAGCCGTACGCAATATC  
 TTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCT  
 ATCAAGCGGTGCGAGATGGGCGCGATCGCATTTGCCGAAAGGGAATCCGCCGTACGCGACAATTTTGGCGATGCGGCA  
 TACGCCAAATACCCGTCCTTACCATTTCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAAGAAAAACATC  
 ACCTCCTCAACCGTGCCGCCGTCAAACGGCAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCG  
 TTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCCTGA

1 MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAETK  
 51 YNGDSLNTG KLKNDKVSFR DFIRQIEVDG QLITLESGEF QVYKQSHSAL  
 101 TAFQTEQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA  
 25 151 FGSDDAGGKL TYTIDFAAQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
 201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL  
 251 DGGGGTGSSD LANDSFIRQV LDRQHFBPDG KYHLFGSRGE LAERSGHIGL  
 301 GKIQSHQLGN LMIQQAAIKG NIGYIVRFS D HGHEVHSPFD NHASHSDSDE  
 351 AGSPVDGFS L YRIHWDGYEH HPADGYDGPQ GGGYPAPKGA RDIYSYDIKG  
 30 401 VAQNIRLNL T DNRSTGQRLA DRFHNAGSML TQGVGDGFKR ATRYSPELDR  
 451 SGNAAEAFNG TADIVKNIIG AAGEIVGAGD AVQGISSEGSN IAVMHGLGLL  
 501 STENKMARIN DLADMAQLKD YAAAAIRDWA VQNPNAAQGI EAVSNIFMAA  
 551 IPIKGIGAVR GKYGLGGITA HPIKRSQMG A IALPKGKSAV SDNFADAAYA  
 601 KYPSPYHSRN IRSNLEQRYG KENITSS T V P PSNGKNVKLA DQRHPKTGVP  
 35 651 FDGKGFPNFE KHVKYDTLEH HHHHH\*

### Example 5 – hybrids of 287

Expression of 287 as full-length with a C-terminal His-tag, or without its leader peptide but  
 with a C-terminal His-tag, gives fairly low expression levels. Better expression is achieved  
 40 using a N-terminal GST-fusion. As an alternative to using GST as an N-terminal fusion  
 partner, 287 was placed at the C-terminus of protein 919 ('919-287'), of protein 953 ('953-  
 287'), and of proteins ORF46.1 ('ORF46.1-287'). In both cases, the leader peptides were  
 deleted, and the hybrids were direct in-frame fusions.

To generate the 953-287 hybrid, the leader peptides of the two proteins were omitted by  
 45 designing the forward primer downstream from the leader of each sequence; the stop codon  
 sequence was omitted in the 953 reverse primer but included in the 287 reverse primer. For  
 the 953 gene, the 5' and the 3' primers used for amplification included a *NdeI* and a *BamHI*  
 restriction sites respectively, whereas for the amplification of the 287 gene the 5' and the 3'  
 primers included a *BamHI* and a *XhoI* restriction sites respectively. In this way a sequential  
 50 directional cloning of the two genes in pET21b+, using *NdeI-BamHI* (to clone the first gene)  
 and subsequently *BamHI-XhoI* (to clone the second gene) could be achieved.

The 919-287 hybrid was obtained by cloning the sequence coding for the mature portion of 287 into the *Xho*I site at the 3'-end of the 919-His clone in pET21b+. The primers used for amplification of the 287 gene were designed for introducing a *Sal*I restriction site at the 5'- and a *Xho*I site at the 3'- of the PCR fragment. Since the cohesive ends produced by the *Sal*I and *Xho*I restriction enzymes are compatible, the 287 PCR product digested with *Sal*I-*Xho*I could be inserted in the pET21b-919 clone cleaved with *Xho*I.

The ORF46.1-287 hybrid was obtained similarly.

The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens:

	Mixture with 287	Hybrid with 287
<b>919</b>	32000	16000
<b>953</b>	8192	8192
<b>ORF46.1</b>	128	8192

Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained for 919-287 and 953-287:

	<b>919</b>		<b>953</b>		<b>ORF46.1</b>	
<i>Strain</i>	<i>Mixture</i>	<i>Hybrid</i>	<i>Mixture</i>	<i>Hybrid</i>	<i>Mixture</i>	<i>Hybrid</i>
<b>MC58</b>	512	1024	512	1024	-	1024
<b>NGH38</b>	1024	2048	2048	4096	-	4096
<b>BZ232</b>	512	128	1024	16	-	-
<b>MenA (F6124)</b>	512	2048	2048	32	-	1024
<b>MenC (C11)</b>	>2048	n.d.	>2048	n.d.	-	n.d.
<b>MenC (BZ133)</b>	>4096	>8192	>4096	<16	-	2048

Hybrids of ORF46.1 and 919 were also constructed. Best results (four-fold higher titre) were achieved with 919 at the N-terminus.

Hybrids 919-519His, ORF97-225His and 225-ORF97His were also tested. These gave moderate ELISA titres and bactericidal antibody responses.

As hybrids of two proteins A & B may be either NH<sub>2</sub>-A-B-COOH or NH<sub>2</sub>-B-A-COOH, the "reverse" hybrids with 287 at the N-terminus were also made, but using ΔG287. A panel of strains was used, including homologous strain 2996. FCA was used as adjuvant:



	287 & 919		287 & 953		287 & ORF46.1	
Strain	$\Delta G287-919$	919-287	$\Delta G287-953$	953-287	$\Delta G287-46.1$	46.1-287
2996	128000	16000	65536	8192	16384	8192
BZ232	256	128	128	<4	<4	<4
1000	2048	<4	<4	<4	<4	<4
MC58	8192	1024	16384	1024	512	128
NGH38	32000	2048	>2048	4096	16384	4096
394/98	4096	32	256	128	128	16
MenA (F6124)	32000	2048	>2048	32	8192	1024
MenC (BZ133)	64000	>8192	>8192	<16	8192	2048

Better bactericidal titres are generally seen with 287 at the N-terminus.

When fused to protein 961 [ $\text{NH}_2$ - $\Delta G287$ -961-COOH – sequence shown above], the resulting protein is insoluble and must be denatured and renatured for purification. Following renaturation, around 50% of the protein was found to remain insoluble. The soluble and insoluble proteins were compared, and much better bactericidal titres were obtained with the soluble protein (FCA as adjuvant):

	2996	BZ232	MC58	NGH38	F6124	BZ133
Soluble	65536	128	4096	>2048	>2048	4096
Insoluble	8192	<4	<4	16	n.d.	n.d.

Titres with the insoluble form were, however, improved by using alum adjuvant instead:

Insoluble	32768	128	4096	>2048	>2048	2048
-----------	-------	-----	------	-------	-------	------

- 961c was also used in hybrid proteins (see above). As 961 and its domain variants direct efficient expression, they are ideally suited as the N-terminal portion of a hybrid protein.

### Example 23 – further hybrids

Further hybrid proteins of the invention are shown in the drawings and have the sequences set out below. These are advantageous when compared to the individual proteins:

#### ORF46.1-741

ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTTGAACCCGACGGGAAATACCAC  
CTATTCGGCAGCAGGGGGGAACCTTGCCGAGCGCAGCGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC  
AACCTGATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTC  
CATTCCTCCCTTCGACAACCATGCTTACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTAC  
CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGCTATCCCGCTCCC  
AAAGGCGCAGGGATATATACAGCTACGACATAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGC  
AGCACCGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA  
CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTT  
AAAAACATCATCGGCGCGGCAGGAGAAATTGTTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAACATT

GCTGTCATGCACGGCTTGGGTCTGCTTTCCACCGAAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA  
 CTCAAAGACTATGCCGAGCAGCCATCCCGGATTGGGCGAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC  
 AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCAGC  
 5 GCACATCCTATCAAGCGGTCGAGATGGGCGCGATCGCATTTGCCGAAAGGGAAATCCGCCCTCAGCGACAATTTTGCC  
 GATGCGGCATACGCCAAATACCCGTCCTTACCATTCCCGAAATATCCGTTCAAACCTTGGAGCAGCGTTACGGCAA  
 GAAAACATCACCTCCTCAACCGTGCCGCGTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACA  
 GCGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGGGGTGGT  
 10 GTCGCGCGGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGAGTCT  
 TTGACGCTGGATCAGTCCGTGAGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAC  
 GGTGACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTGAGCGTTTCGACTTTATCCGCCAAATCGAAGTGGAC  
 GGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCCTTAACCGCCTTTCAGACC  
 GAGCAAAATACAGATTCCGAGCATTCGGGAAGATGTTGCGAAACGCCAGTTCAGAACTCGGCGACATAGCGGGCGAA  
 CATACATCTTTTGACAAGCTTCCCGAAGCGCGGACGGCGACATATCGCGGGACGGCGTTCGAGTTCAGACGATGCCGCG  
 15 GGAACCTGACCTACACCATAGATTTGCGCGCAAGCAGGGAACCGGCAAAATCGAACATTTGAAATCGCCAGAACTC  
 AATGTCGACCTGGCCGCGCGCATATCAAGCCGATGGAACCGCATGCCGTCATCAGCGGTTCCGTCTTTACAAC  
 CAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGCGAGCGCGGAAGTGAA  
 ACCGTAAACGGCATACGCCATATCGGCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCTGA

1 MSDLANDSFI RQVLDRQHF PDGKYHLFGS RGE LAERSGH IGLGKIQSHQ  
 51 LGNLMIQQA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG  
 101 FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL  
 151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRA TRYSP E LDRSGNAABEA  
 201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA  
 251 RINDLADMAQ LKDYAAAAIR DWAVQNPNA QGIEAVSNIF MAAPIKIGI  
 25 301 AVRKYGLGG ITAHPKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH  
 351 SRNIRSNLEQ RYKKNITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGFP  
 401 NFEKHVKYDT GSGGGVAAAD IGAGLADALT APLDHKDKGL QSLTLDQSVR  
 451 KNEKLKLAQ GAEKTYGNGD SLNTGKLKND KVS RFD FIRQ IEVDGQLITL  
 501 ESGEFQVYKQ SHSALTAFQT EQIQDSEHS KMVAKRQFRI GDIAGEHTSF  
 30 551 DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAKQNGKRI EHLKSPELNV  
 601 DLAAADIKPD GKRHAIVSGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK  
 651 TVNGIRHIGL AAKQLEHHHH HH\*

35 **ORF46.1-961**  
 ATGTCAGATTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTGCAACCCGACGGGAAATACCAC  
 CTATTCCGGCAGCAGGGGGAACTTGCCGAGCGCAGCGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC  
 AACCTGATGATTCAACAGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCAGGGGCAGGAAGTC  
 40 CATTCCTCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTAC  
 CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGCGGCTATCCCGCTCCC  
 AAAGGCGCGAGGGATATATACAGCTACGACATAAAAGCGCTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGC  
 AGCACCGGACACGGCTTGCCGACCGTTCCACAATGCCGCTAGTAGCTGACGCAAGTAGGCGCAGGATTCAAA  
 CGCGCCACCCGATACAGCCCGAGCTGGACAGATCGGGCAATGCCGCCAAGCCTTCAACCGCACTGACAGATATCGTT  
 45 AAAAAATCATCGGCGCGGAGGAGAAATGTGCGCGCAGGCGATGCCGTGACGGGCATAAGCGAAGGCTCAAACATT  
 GCTGTCATGCACGGCTTGGGTCTGCTTTCCACCGAAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA  
 CTCAAAGACTATGCCGAGCAGCCATCCCGGATTGGGCGAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC  
 AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCAGC  
 GCACATCTATCAAGCGGTGCGAGATGGGCGCGATCGCATTTGCCGAAAGGAAATCCCGCTCAGCACAATTTTGCC  
 50 GATGCGGCATACGCCAAATACCCGTCCTTACCATTCCCGAAATATCCGTTCAAACCTTGGAGCAGCGTTACGGCAA  
 GAAAACATCACCTCCTCAACCGTGCCGCGTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACA  
 GCGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGAGGAGGA  
 GCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAGAAATCAAC  
 GTTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGAT  
 55 GTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAACAAA  
 CAAAACGTCGATGCCAAAGTAAAGCTGCGAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCC  
 GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTT  
 GCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCC  
 GAAGCATTCAACGATATCGCGGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAA  
 60 GCCAAACAGACGGCCGAAGAAACCAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGGCC  
 GAAGCTGCGCTGGCACAGCTAATACTGCAGCCGACAAAGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCT  
 GATATCGCTACGAACAAAGATAATTTGCTAAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGC  
 AAATTTGTGAGAATTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTGGCTTCTGCTGAAAAATCCATT  
 65 GCCGATCACGATACTCGCTGAACGGTTTGGATAAAACAGTGTCAGACCTGCGCAAAAGAAACCCGCCAAGGCCCTTGCA  
 GAACAAGCCGCGCTCTCCGGTCTGTTTCAACCTTACAACGTGGGTGCGTTCAATGTAACGGCTGCAGTCGCGCGGTAC  
 AAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAGCAGGCGTGGCAGTC  
 GGCACATTCTCGGTTCTTCCGACGCTACCATGTGCGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCTGA

1 MSDLANDSFI RQVLDROHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ  
 51 LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG  
 101 FSLYRIHWDG YEHPADGYD GPQGGYPAP KGARDIYSYD IKGVAQNIRL  
 151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSP E LDRSGNAAEA  
 201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA  
 251 RINDLADMAQ LKDYAAAAIR DWAVQNPNA QGIEAVSNIF MAAIPIKGIG  
 301 AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH  
 351 SRNIRSNLEQ RYKGENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGF  
 401 NFEKHVKYDT GSGGGGATND DDVKAATVA IAAAYNNGQE INGFKAGETI  
 451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK  
 501 VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLK ENITTFAEET  
 551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK  
 601 QTAETKQNV DAKVKAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA  
 651 DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK  
 701 SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFPQY NVGRFNVTA  
 751 VGGYKSESAV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY HVGVNYEWLE  
 801 HHHHHH\*

**ORF46.1-961c**

ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCAC  
 CTATTCGGCAGCAGGGGGGAACCTTCCGAGCGCAGCGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGC  
 AACCTGATGATTCACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCAGGGCAGCAAGTC  
 CATTCCCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTAC  
 CGCATCCATTGGGACGGATACGAACACCATCCCGCGCAGCGGTATGACGGGCCACAGGGCGGCGGTATCCCGTCCCC  
 AAAGGCGCGAGGATATATACAGCTACGACATAAAAGCGGTTGCCCAAATATCCGCTCAACCTGACCGACAACCGC  
 AGCACCGGACAACGCTTGGCGACCGCTTCCACAATGCCGCTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA  
 CGCGCCACCCGATACAGCCCGAGCTGGACAGATCGGGCAATGCCGCGGAAGCCTTCAACGGCACTGCAGATATCGTT  
 AAAACATCATCGGCGCGGAGAGAAATGTGCGCGCAGGCGATGCCGTGACGGGCATAAGCGAAGGCTCAAACATT  
 GCTGTATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTGGCAGATATGGCGCAA  
 CTCAAAGACTATGCCGAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC  
 AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATTGGAGCTGTTGGGGAAAAATACGGCTTGGGCGGCATCAG  
 GCACATCTATCAAGCGGTCGAGATGGGCGCGATCGCATTCGCGAAAGGAAATCCGCGCTAGCGACAATTTGCC  
 GATGCGGCATACGCCAAATACCCGTCCTTACCATTCCCGAAATATCCGTTCAAACCTGGAGCAGCACTGATGCC  
 GAAACATCACCTCTCAACCGTGGCGCGCTCAAACGGCAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACA  
 GCGGTACCGTTTGACGGTAAAGGGTTTCCGAATTTTGAAGAAGCAGCTGAAATATGATACGGGATCCGGAGGAGGAGGA  
 GCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCATAACAATGGCCAAAGAAATCAAC  
 GGTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGAT  
 GTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAACAAA  
 CAAAACGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCC  
 GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTGAATAAATTGGGAGAAAAATATAACGACATTT  
 GCTGAAGAGACTAAGACAAATATCGTAAAAAATTGATGAAAAATTAGAAGCCGTGCTGATACCGTGCACAAGCATGCC  
 GAAGCATCTCAACGATATCGCGATTTCATTGGATGAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAA  
 GCCAAACAGCAGCGCCGAAGAAACCAACGCTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGGCC  
 GAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAGGCGCAAGCTGTCGCTGCAAAAGTTACCGACATCAAAGCT  
 GATATCGCTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGCTGACACAGGAGAGTCTGCACAGC  
 AAATTTGTGAGAATTGATGCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATT  
 GCCGATCACGATACTCGCTGAACGGTTTGGATAAAACAGTGTACAGCCTGCGCAAAGAAACCGCCAAGGCCTTGCA  
 GAACAAGCCGCGCTCTCCGGTCTGTTCACCTTACAACGTGGGTCTCGAGCACCACCACCACCACCTGA

1 MSDLANDSFI RQVLDROHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ  
 51 LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG  
 101 FSLYRIHWDG YEHPADGYD GPQGGYPAP KGARDIYSYD IKGVAQNIRL  
 151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSP E LDRSGNAAEA  
 201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA  
 251 RINDLADMAQ LKDYAAAAIR DWAVQNPNA QGIEAVSNIF MAAIPIKGIG  
 301 AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH  
 351 SRNIRSNLEQ RYKGENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGF  
 401 NFEKHVKYDT GSGGGGATND DDVKAATVA IAAAYNNGQE INGFKAGETI  
 451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK  
 501 VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLK ENITTFAEET  
 551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK  
 601 QTAETKQNV DAKVKAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA  
 651 DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK  
 701 SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFPQY NVGLEHHHHH  
 751 H\*

**961-ORF46.1**

5 ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC  
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
 10 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAAC  
 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT  
 GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTTGGGAGAAAAATATAACGACA  
 TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
 15 GCCGAAGCATTCAACGATATCGCCGATTCAATTGGATGAAACCAACACTAAGGCAGACGAAGCCCGTCAAAACCGCCCAAT  
 GAAGCCAAACAGACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCAAA  
 GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA  
 CTTGATATCGCTACGAACAAAGATAATATGCTAAAAAAGCAAACAGTGCCGACGTGTACACCCAGAGAAGAGTCTGAC  
 AGCAAAATTTGTCAGAATTGATGGTCTGAAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC  
 20 ATTGCCGATCAGGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAGAAACCCGCCAAGGCCCTT  
 GCAGAACAAAGCCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCTGGTTCAATGTAACGGCTGCAGTCGGCGGC  
 TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA  
 GTCGGCACTTCGTCCGGTCTTCCGCAGCCTACCATGTCCGGCTCAATTACGAGTGGGGATCCGGAGGAGGAGGATCA  
 GATTTGGCAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTTC  
 25 GGCAGCAGGGGGAACTTGGCAGCGCAGCGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTG  
 ATGATTCAACAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCAGGGGCACGAAGTCCATTCC  
 CCCCTTCGACAACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCCTTTACCGCATC  
 CATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGCGGTATCCCGCTCCCAAAGGC  
 GCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCCGACGACC  
 30 GGACAACGGCTTGGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAACACGCGCC  
 ACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAAC  
 ATCATCGGCGCGGAGGAGAAATTTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTC  
 ATGCACGGCTTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTGGCAGATATGGCGCAACTCAAA  
 GACTATGCCCGCAGCAGCCATCCCGGATTGGGCGAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTGAGCAAT  
 35 ATCTTTATGGCAGCCATCCCCATCAAAGGATTGGAGCTGTTCCGGGAAAAATACGGCTTGGGCGGCATCAGGCGACAT  
 CCTATCAAGCGGTGCGAGATGGGCGCGATCGCATTGCCGAAAGGGAAATCCGCGCTGACGACACAATTTTGCCGATGCG  
 GCATACGCCAAATACCCGTCCCTTACCATTCCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAGAGAAAC  
 ATCACCTCCTCAACCGTGCCGCGCTCAAACGGCAAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTA  
 CCGTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCAC  
 TGA

1 MATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDFK GLGLKVVVN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL  
 101 ADTDAALADT DAALDATTNA LNKLGENTIT FAEETKTNIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEEVAKV TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 301 VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK SESAVAIGTG  
 351 FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEWGSGGGGS DLANDSFIRQ  
 401 VLDRQHFEFD GKYHLFSGR ELAERSGHIG LGKIQSHQLG NLMIQQAIAK  
 451 GNIGYIVRFS DHGHEVHSPF DNHASHSDSD EAGSPVDGFS LYRIHWDGYE  
 501 HHPADGYDGP QGGGYPAKPG ARDIYSYDIK GVAQNIRLNL TDNRSTGQRL  
 551 ADRFHNAGSM LTQVGVDGFK RATRYSPELD RSGNAAEAFN GTADIVKNII  
 601 GAAGEIVGAG DAVQGISEGS NIAVMHGLGL LSTENKMARI NDLADMAQLK  
 651 DYAAAIRDW AVQNPNAAG IEAVSNIFMA AIPKIGIGAV RGKYGLGGIT  
 701 AHPIKRSQMG AIALPKGKSA VSDNFADAAY AKYPSFYHSR NIRSLEQRY  
 751 GKENITSSTV PPSNGKNVKL ADQRHPKTGV PFDGKGFPNF EKHVKYDTLE  
 801 HHHHHH\*

**961-741**

55 ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC  
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAAC  
 60 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT  
 TCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCAACCGCTTGAATAAATTTGGGAGAAAAATATAACGACA  
 GTCGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
 GCCGAAGCATTCAACGATATCGCCGATTCAATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT  
 GAAGCCAAACAGACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCAAA  
 65 GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAA  
 GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCCGACGTGTACACCCAGAGAAGAGTCTGAC  
 AGCAAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC  
 ATTGCCGATCAGGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACAGCTGCGCAAGAAACCCGCCAAGGCCCTT

GCAGAACAAAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGCGTTCAATGTAACGGCTGCAGTCGGCGGC  
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA  
5 GTCCGCACTTCGTCCGGTCTTCCGCGAGCTTACCATGTCGGCGTCAATTACGAGTGGGGATCCGGAGGGGGTGGTGTG  
GCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGAGTCTTTG  
ACGCTGGATCAGTCCGTGAGGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAACCGGT  
GACAGCCTCAATACGGGCAAATTGAAGAACGACAAGGTGAGCGGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGG  
CAGCTCATTACCTTGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCTTTAGACCGAG  
10 CAAATACAAGATTTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTTCAAGTTCGGCGACATAGCGGGCGAACAT  
ACATCTTTTGACAAGCTTCCCGAAGGCGGAGGCGACATATCGCGGACGGCGTTCGGTTTACAGACGATGCCGGCGGA  
AACTGACCTACACCATAGATTTCGCCGCCAAGCAGGGAACGGCAAAATCGAACATTGAAATCGCCAGAACTCAAT  
GTCGACCTGGCGCGCGGATATCAAGCCGATGGAACCGCATGCCGTATCAGCGGTTCCGCTCTTACAACCAA  
GCCGAGAAAGGCGATTACTCCCTCGGTATCTTTGGCGGAAACGCCAGGAAGTTGCGCGCAGCGGAAGTGAAGAAC  
GTAACGGCATAACGCATATCGGCCTTGCCGCCAAGCAACTCGAGCACCACCACCACCACCTGA

15 1 MATNDDVVK AATVAIAAAY NNGQINGFK AGETIYDIDE DGTITKDAT  
51 AADVEADDFK GLGLKVVVN LTKTVNENKQ NVDKVKAAE SEIEKLTKL  
101 AADDAALADT DAALDATTNA LNKLGENTT FAEETKTIV KIDEKLEAVA  
151 DTVDKHABAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAVK  
201 AAETAAGKAE AAAGTANTAA DKAEEVAKV TDIKADIATN KDNIAKKANS  
20 ADVYTRESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
301 VSDLRKETRQ GLAEQAALSG LFPYNVGRF NVTAAVGGYK SESAVAIGTG  
351 FRFTENFAAK AGVAVGTSSG SSAAYHVGUN YEWGSGGGV AADIGAGLAD  
401 ALTAPLDHK KGLQSLTDQ SVRKNEKLK AAQGAERTYG NGDSLNTGKL  
451 KNDKVSFRDF IRQIEVDGQL ITLESGEFQV YKQSHSALTA FQTEQIQDSE  
25 501 HSGKMAVRQ FRIGDIAGEH TSFDKLPEGG RATYRGTAFG SDDAGGKLT  
551 TIDFAAQGN GKIEHLKSP LNVDLAAADI KPDGKRHAVI SGSVLYNQAE  
601 KGSYSLGIFG GKAEVAGSA EVKTVNGIRH IGLAAKQLEH HHHHH\*

30 **961-983**  
ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAAGAAATC  
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
35 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAA  
AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAGTTAAACCAAGTTAGCAGACACTGAT  
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCAACGCCCTTGAATAAATTGGGAGAAAATATAACGACA  
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
GCCGAAGCATTCAACGATATCGCCGATTCTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCCGCCAAT  
40 GAAGCCAAACAGACGGCCGAAGAAACCAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAA  
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAA  
GCTGATATCGTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGTGTACACGAGAGAAGAGTCTGAC  
AGCAAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC  
ATTGCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAGAAACCCGCCAAGGCCCTT  
45 GCAGAACAAAGCCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGCGTTCAATGTAACGGCTGCAGTCGGCGGC  
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCA  
GTCGGCACTTCGTCCGGTCTTCCGCGAGCTTACCATGTCGGCGTCAATTACGAGTGGGGATCCGGCGAGGCGGCACT  
TCTGCGCCGCACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACACAGGAAATCAGCAGCAGTA  
50 TCTTACCGCGGTATCAAGAACGAAATGTGCAAGACAGCATGCTCTGTGCGCGGTGCGGATCGGCTGCGGATACG  
GACAGGGATGCCAAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCAATGACGCATACAAG  
AATTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGTAGAGGTAGGTATCGTCGACACAGGC  
55 GAATCCGTCGGCAGCATATCCTTTCCCGAATGTATGGCAGAAAAGAACCGGTATAACGAAAATTACAAAACATAT  
ACGGCGTATATGCGGAAGGAAGCGCTGAAGACGAGGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAGGCCGTT  
ATAGAGACTGAAGCAAAGCCGACGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCATATTATT  
GGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGAT  
60 GAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAAGCTGGCGTGGCATCGTC  
AATAACAGTTTTTGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTTCGGAGGAGCAGTACCGC  
CAAGCGTTGCTCGACTATTCCGCGCGGTGATAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGCAAC  
CTGTCTTACCACATCCGTAATAAAAAATGCTTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACA  
75 TATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGACGGCTAGACCGCAGTGGAGAA  
AAGTTCAACAGGGAATGTATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCAACCATTCGGAATTACTGCC  
ATGTGGTGCTGTGCGCACCCATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTGCCGGAACATCC  
TTTTCGCGACCCATCGTAACCGGCACGGCGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGT  
ACCAAGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCCGCTGGGGACTGCTGGATGCG  
GGTAAGGCCATGAACGACCGCGCTCTTCCGTTCCGGCGACTTTACCGCGGATACGAAAGGTACATCCGATATTGCC  
80 TACTCCTTCCGTAACGACATTTTCAGGACCGGCGGCTTCGATCAAAAAGGCGGCGAGCAACTGCAACTGCACGGCAAC  
AACACCTATACGGGCAAAACCATTATCGAAGCGGCTGCTGTTGTACGGCAACAAACAAATACGATACGCGGTC  
GAAACCAAGGTGCGCTGATTATAACGGGCGGCATCCGGCGGAGCCTGAACAGCGACGGCATTTGCTATCTGGCA  
GATACCGACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACGCTGTAC

5 ACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTCGGCACGCGGCAAG  
 GGGGACAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTC  
 TTCACAAACATCGAAACCGACGGCGGCTGCTGGCTTCCTTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCGAC  
 10 ACGTGTCTTATTATGTCCGTCGCGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCGCGCCCGCCGGTCTG  
 AAACACGCCGTAGAACAGGGCGGCAGCAATCTGGAAACCTGATGGTTCGAACCTGGATGCCTCCGAATCATCCGCAACA  
 CCGGAGACGGTTGAAACTGCGGCAGCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTCCGCGCA  
 GCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGTACCGTCTATGCCGAC  
 AGTACCGCCGCCCATGCCGATATGCAGGGACGCCGCTGAAAGCCGTATCGGACGGGTGGACCACAACGGCAGCGGT  
 15 CTGCGCGTCATCGCGCAACCCAACAGGACGGTGGAACTGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGCAGT  
 ACCCAAACCGTCGGCATTGCCCGCAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGACGACACA  
 TGGAGCGAAAACAGTGCAAAATGCAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGCATGCGGGCGATATC  
 GGCATATCTCAAAGGCCCTGTTCTCTACGGACGCTACAAAAACAGCATACGCCGACAGCAGCGGTGCGGACGAACATGCG  
 GAAGGCAGCGTCAACGGCAGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGAT  
 20 TTGACGGTTCGAAGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTGCGCGAAAAAGGCAGTGCTTTGGGCTGG  
 AGCGGCAACAGCCTCACTGAAGGCACGCTGGTGGACTCGCGGTCTGAAGCTGTCGCAACCCCTTGAGCGATAAAGCC  
 GTCCGTTTGGCAACGGCGGGCGGTGGAACGCGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACCGGCGCG  
 ACTGCGCAACCGGCAAGACGGGGGCACGCAATATGCCGCACACCCGCTCTGTTGCGCGCTGGGCGCGGATGTCGAA  
 TTCGGCAACGGCTGGAACGGCTTGGCAGTTACAGCTACGCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTC  
 GCGGTAGGCTACCGGTTCTCGAGCACCACCACCACCACCTGA

20

1 MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL  
 101 ADTDAALADT DAALDATTNA LNKLGENTIT FAETRTNIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEB TKQNVDAKVK  
 25 201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 301 VSDLRKETRQ GLAEQAALSG LFQPINVGRF NVTAAVGGYK SESAVAIGTG  
 351 FRFTENFAAK AGVAVGTSSG SSAAYHVGUN YEWGSGGGGT SAPDFNAGGT  
 401 GIGSNSRATT AKSAAVSYAG IKNEMCKDRS MLCAGRDDVA VTDRDAKINA  
 30 451 PPPNLHTGDF PNPNDAYKNL INLKPAIEAG YTGRGVEVGI VDTGESVSGSI  
 501 SFPELYGRKE HGVNENYKNY TAYMRKEAPE DGGGKDIEAS FDDEAVIETE  
 551 AKPTDIRHVK EIGHIDLVS H IIGRSVDGR PAGGIAPDAT LHIMNTNDET  
 601 KNEMMVAAIR NAWVKLGERG VRIVNNSFGT TSRAGTADLF QIANSEEQYR  
 651 QALLDYSGGD KTDEGIRLMQ QSDYGNLSYH IRNKNMLFIF STGNDQAQAP  
 35 701 NITYALLPFYE KDAQKGIITV AGVDRSGEKF KREMYGEPGT EPLEYGSNHC  
 751 GITAMWCLSA PYEASVRFTR TNPIQIAGTS FSAPIVTGTA ALLLQKYPWM  
 801 SNDNLRITLL TTAQDIGAVG VDSKFGWLL DAGKAMNGPA SFPFGDTAD  
 851 TKGTSDIAYS FRNDISGTGG LIKKGGSQQLQ LHGNNTYTGK TIEGGSVLV  
 901 YGNNSKDMRV ETKGALIYNG AASGGSLSND GIVYLADTDQ SGANETVHIK  
 40 951 GSLQLDGKGT LYTRLGKLLK VDGTAIIGK LYMSARGKA GYLNSTGRRV  
 1001 PFLSAAKIGQ DYSFTNIET DGGLLASLDS VEKTAGSEGD TLSYVVRGN  
 1051 AARTASAAAH SAPAGLKHAV EQGGSNLENL MVELDASESS ATPETVETAA  
 1101 ADRTDMPGIR PYGATFRAAA AVQHANAADG VRIFNSLAAT VYADSTAHA  
 1151 DMQGRRLKAV SDGLDHNGTG LRVIAQTQOD GGTWEQGGVE GKMRGSTQTV  
 45 1201 GIAAKTGENT TAAATLGMGR STWSSENSANA KTDISISLFA IRHDAGDIGY  
 1251 LKGLFSYGRY KNSISRSTGA DEHAEGSVNG TLMQLGALGG VNPFFAATGD  
 1301 LTVEGGLRYD LLKQDAFAEK GSALGWSGNS LTEGLTVGLA GLKLSQPLSD  
 1351 KAVLFATAGV ERDLNGRDYT VTGGFTGATA ATGKTGARNM PHTRLVAGLG  
 50 1401 ADVEFGNGWN GLARYSYAGS KQYGNHSGRV GVGYRFLEHH HHHH\*

50

**961c-ORF46.1**

55 ATGGCCACAAACGACGACGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC  
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAC  
 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAAACAACCAAGTTAGCAGACACTGAT  
 GCGGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTTGGGAGAAAATATAACGACA  
 TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
 60 CCGGAGCATCTCAACGATATCGCCGATTCTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT  
 GAAGCCAAACAGACGGCCGAAGAAACCAAAACAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAA  
 GCCGAAGCTGCCGTGGCAGCTAATACTGCAGCCGACAAGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAAA  
 GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC  
 AGCAAATTTGTCAGAATTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTGGCTTCTGCTGAAAAATCC  
 65 ATTGCCGATCAGGATACGCTGAAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAGAAACCGCCAAAGCTTT  
 CAGAAACAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGAGGAGGATCAGATTTGGCA  
 AACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTTCGAACCCGACGGGAAATACCACCTATTCGGCAGCAGG  
 GGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAA

5 CAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGAC  
 AACCATGCCTCACATTCCGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTACCGCATCCATTGGGAC  
 GGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGCTATCCCGCTCCCAAAGGCGCGAGGGAT  
 10 ATATACGACTACGACATAAAAGGCGTTGCCAAAATATCCGCTCAACCTGACCGACAACCCGAGCACCAGACAACGG  
 CTGCGGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCGCCACCCGATAC  
 AGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGC  
 GCGGCAGGAGAAATTGTGCGCGCAGGCGATGCCGTGACGGGCATAAGCGAAGGCTCAAACATTGCTGTCTATGCACGGC  
 TTGGGTCTGCTTTCCACCAGAAAACAAGATGGCGCGCATCAACGATTGTCAGATATGGCGCAACTCAAAGACTATGCC  
 15 GCAGCAGCCATCCCGGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATG  
 GCAGCCATCCCCATCAAAGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCTATCAAG  
 CGGTCGCAGATGGGCGCATCGCATTGCCGAAAGGAAATCCGCGTCAGCGACAATTTTGGCGATGCGGCATACCGCC  
 AAATACCCGTCCTTTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAGAAAACATCACCTCC  
 TCAACCGTGCCGCGCTCAAACGGCAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGAC  
 20 GGTAAAGGTTTCCGAATTTTGAAGACAGTGAAATATGATACGCTCGAGCACCACCACCACCACCTGA

1 MATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDKVKAAE SEIEKLTKL  
 101 ADTDAALADT DAALDATNA LNKLGENTT FAEETKTIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 301 VSDLRKETRO GLAEQAALSG LFQPYNVGGS GGGGSDLAND SFIRQVLDRO  
 351 HFEPDGKYLH FGSRGELAER SGHIGLGIQ SHQLGNLMIQ QAAIKGNIGY  
 401 IVRFSDHGE VHSPPDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHPAD  
 25 451 GYDGPQGGY PAPKGARDY SYDIKVAQN IRLNLTNRS TQRLADRFH  
 501 NAGSMLTQGV GDGFKRATRY SPELDRSNA AEFNGTADI VKNIIAAGE  
 551 IVGAGDAVQG ISEGSNIAMV HGLGLLSTEN KMARINDLAD MAQLKDYAAA  
 601 AIRDWAVQNP NAAQIEAVS NIFMAAPIK GIGAVRGKYG LGGITAHPIK  
 651 RSQMGALALP KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LBQRYGKENI  
 30 701 TSSTVPPSNG KNVKLADQRH PKTGVPFDGK GPFNFKEHVK YDTLEHHHHH  
 751 H\*

35 **961c-741**  
 ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACATGGCCAAGAAATC  
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCTGTGACTAACCTGACCAAAACCGTCAATGAAAAC  
 40 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAAACAACCAAGTTAGCAGACACTGAT  
 CCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACA  
 TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
 GCCGAAGCATTCAACGATATCGCCGATTCTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCCGCCAAT  
 GAAGCCAAACAGACGGCCGAAGAAACCAACAAACAGCTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAA  
 45 GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAA  
 GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC  
 AGCAAAATTTGTCAGAAATGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCC  
 ATTGCCCATACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTACAGCTGCGCAAGAAACCCGCCAAGGCCTT  
 GCAGAACAAAGCCGCTCTCCGCTGTGTTCAACCTTACAACCTGGGTGGATCCGGAGGGGGTGGTGTCCGCCCGGAC  
 ATCGGTGCGGGGCTTCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGAT  
 50 CAGTCCGTGAGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAATTATGGAACGGTGACAGCCTC  
 AATACGGGCAATTTGAAGAACGACAAGGTGAGCCGTTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCATT  
 ACCTTGGAGAGTGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCTTAAACCGCTTTTCAGACCGAGCAATACAA  
 GATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTTCAAGATCGGCGACATAGCGGGCGAACATACATCTTTT  
 GACAAGCTTCCGGAAGCGGCAGGCGACATATCGCGGGACGCGTTCCGGTTCAGACGATGCCGGCGGAAAACTGACC  
 55 TACACCATAGATTTCCGCGCAAGCAGGGAACGGCAAAATCGAACATTGAAATCGCCAGAACTCAATGTGACCTG  
 GCCGCCCGGATATCAAGCCGGATGGAACCGCATGCCGTATCAGCGGTTCCGTCTTTACAACCAAGCCGAGAAA  
 GGCAGTTACTCCCTCGGTATCTTTGGCGGAAAGCCAGGAAGTTGCCGGCAGCGCGGAAGTGAACCCGTAAACGGC  
 ATACGCCATATCGGCCTTGCCGCAAGCAACTCGAGCACCACCACCACCACCTGA

60 1 MATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDKVKAAE SEIEKLTKL  
 101 ADTDAALADT DAALDATNA LNKLGENTT FAEETKTIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 65 301 VSDLRKETRO GLAEQAALSG LFQPYNVGGS GGGVADIG AGLADALTAP  
 351 LDHKDKGLQS LTLDQSVRKN EKLKLAQGA EPTYGNGDSL NTGKLKNDKV  
 401 SRFDFIRQIE VDGQLITLES GEFQVYKQSH SALTAFTQEQ IQDSEHSGKM



-31-

451 VAKRQFRIGD IAGEHTSFDK LPEGGRATYR GTAFGSDDAG GKLTYTIDFA  
 501 AKQGNKIEH LKSPELNVDL AAADIKPDGK RHAVISGSVL YNQAERKSYS  
 551 LGIFGGKAQE VAGSAEVKTV NGIRHIGLAA KQLEHHHHHH \*

5

**961c-983**

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC  
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC  
 10 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAC  
 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT  
 GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTGGGAGAAAAATATAACGACA  
 TTTGTCTGAAGAGACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT  
 15 GCCGAAGCATTTCAACGATATCGCCGATTCAATTGGATGAAACCAACACTAAGGCAGACAGAACGCCGTCAAAACCGCCAAT  
 GAAGCCAAACAGACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCCAA  
 GCCGAAGCTGCCGCTGGCAGACTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAA  
 GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGAC  
 AGCAAAATTGTGAGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGCTGCTGAAAAATCC  
 ATTGCCGATCAGGATACCTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAGAAAGAACCCGCCAAGGCCTT  
 20 GCAGAACAAGCCGCGCTCTCCGGTCTGTTCACCTTTACAACGTGGGTGGATCCGGCGAGGCGGCACCTTCTCGAGCC  
 GACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCC  
 GGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGGTCCGGATGACGTTGCGGTTACAGACAGGGAT  
 GCCAAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGATACAGAATTTGATC  
 AACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACCGCGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTC  
 25 AGCCTCATCTCTTTCCGAAGCTGTATGGCAGAAAAGAACACGGCTATAACGAAATTTACAAAACTATACGGCGTAT  
 ATGCGGAAGGAAGCGCCTGAAGACGGAGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAGGCGGTATGACGACT  
 GAAGCAAAGCCGACGGATATCCGCCACGTAAAGAAATCGGACACATCGATTGGTCTCCCATATTATTATGCGGGCGT  
 TCCGTGGACGGCAGACCTGCAGGCGGTATTGCGCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAG  
 AACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTGGCGTGCAGCATCGTCAATAACAGT  
 30 TTTGGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTTCGGAGGAGCAGTACCGCCAAGCGTTG  
 CACATCCGTAATAAAAAACATGCTTTTTCATCTTTTCGACAGGCAATGACGCAAGCTCAGCCCAACACATATGCCCTA  
 TTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCGCAGTGGAGAAAAGTTCAAA  
 CGGGAAATGTATGGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGCGGAATTACTGCCATGTGGTGC  
 35 CTGTGCGCACCTTATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTTGCCGAACATCTTTTCCGCA  
 CCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTG  
 CTGACGACGGCTCAGGACATCGGTGCGAGTCGGCGTGAGCAGCAAGTTCCGGCTGGGGACTGCTGGATGCGGGTAAGGCC  
 ATGAACGGACCCGCGTCTTTCCGTTCCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCTACTCCTTC  
 CGTAACGACATTTTCAGGCACGGCGGCTGATCAAAAAGGCGGCAGCCAACCTGCAACTGCACGGCAACAACACCTAT  
 40 ACGGGCAAAACCATTTATCGAAGGCGGTTGCTGGTGTGTACGGCAACAACAATCGGATATGCGCGTGAACACCAAA  
 GGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCGACGGCATTTGTCTATCTGGCAGATACCGAC  
 CAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACGCTGTACACAGCTTTG  
 45 GGCAAACTGTCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTGCGGCAAGGGGCGAGGC  
 TATCTCAACAGTACCGGACGACGTGTTCCTTCTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTTCAAAAAC  
 ATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGGACACGCTGTCC  
 50 TATTATGTCCGTGCGGCAATGCGGACGGACTGCTTCGGCAGCGGCACATTCGCGCGCCGCGGCTGAAACACGCC  
 GTAGAACAGGGCGGCAGCAATCTGAAAAACCTGATGGTGAACCTGGATGCCCTCCGAATCATCCGCAACACCCGAGACG  
 GTTGAACCTGCGGCGAGCGACCGTACAGATATCCGGGCACTCCGCCCCACGGCGCAACTTTCCGCGCAGCGGCAGCC  
 GTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGCTATGCGGACAGTACCGCC  
 55 GCCCATGCGGATATGCAGGACGCGGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGCAGCGGTCTGCGCGTC  
 ATCGCGCAAAACCAACAGGACGGTGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAATGCGCGGCAAGTACCCAAACC  
 GTCGGCATTTGCCGCAAAACCGGCGAAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGACGACATGGAGCGAA  
 AACAGTGAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGATGCGGGCGATATCGGCTATCTC  
 60 AAAGGCTGTCTCTCTACGGACGCTACAAAAACAGCATCAGCCGACGACCGGTGCGGACGAACATGCGGAAGGCAGC  
 GTCAACGGCAGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACGGGAGATTTGACGGTC  
 55 GAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAAGGCAGTGTCTTGGGCTGGAGCGGCAAC  
 AGCCTCACTGAAGGCACGCTGGTTCGACTCGCGGTCTGAAGCTGTGCAACCCCTTGAGCGATAAAGCCGTCTGTTT  
 GCAACGGCGGGCGTGAACGCGACCTGAACGGACGCGACTACACGGTAACGGCGGCTTTACCGGCGGACTGCAGCA  
 ACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGCTGGTTGCGGCGCTGGGCGCGGATGTGCAATTCGGCAAC  
 GGCTGGAACGGCTTGGCACGTTACAGCTACGCCGTTTCCAAACAGTACGGCAACCACAGCGGACGAGTGGCGTAGGC  
 TACCGGTTCTCGAGCACCACCACCACCACCTGA

1 MATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKDAT  
 51 AADVEADDFK GLGLKKVVTN LTKTVNENKQ NVDKVKAAE SEIEKLPTKL  
 101 ADTDAALADT DAALDATNTA LNKLGENTIT FAETKTIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAEB TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEVAARKV TDIKADIATN KDNIKAKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT

65



-32-

301 VSDLRKETRO GLAEQAALSG LFQPYNVGGS GGGGTSAPDF NAGGTGIGSN  
 351 SRATTAKSAA VSYAGIKNEM CKDRSMLCAG RDDVAVTDRD AKINAPPPNL  
 401 HTGDFPNPND AYKNLINLKP AIEAGYTGRG VEVGIVDTGE SVGSI SFPEL  
 5 451 YGRKEHGYNE NYKNYTAYMR KEAPEDGGGK DIEASFDEA VIETEAQPTD  
 501 IRHVKEIGHI DLVSHIIGR SVDGRPAGGI APDATLHIMN TNDETKNEMM  
 551 VAAIRNAWVK LGERVIRIVN NSFGTTSRAG TADLFQIANS EEQYRQALLD  
 601 YSGGDKTDEG IRLMQQSDYG NLSYHIRNKN MLFIFSTGND AQAQPNFYAL  
 651 LPFYEKDAQK GIITVAGVDR SGEKFKREMY GEPGTEPLEY GSNHCGITAM  
 10 701 WCLSAPEYAS VRFTRTNPIQ IAGTSFSAPI VTGTAALLLQ KYPWMSNDNL  
 751 RTTLLTTAQD IGAUVGDSKF GWGLLDAGKA MNGPASFPFG DFTADTKGTS  
 801 DIAYSFRNDI SGTGGLIKKG GSQQLQHGNN TYTGKTIIEG GSVLYYGNKN  
 851 SDMRVETKGA LIYNGAASGG SLNSDGIIVL ADTDQSGANE TVHIKGSQQL  
 901 DGKGTLYTRL GKLLKVDGTA IIGGKLYMSA RKGAGYLSN TGRRVFFLSA  
 951 AKIGQDYSFF TNIEDGGLL ASLDSVEKTA GSEGDLSY VRRGNAARTA  
 15 1001 SAAAHSAFAG LKHAVERQGS NLENLMVELD ASESSATPET VETAAADRTD  
 1051 MPGIRPYGAT FRAAAAVQHA NAADGVRIFN SLAATVYADS TAAHADMQGR  
 1101 RLKAVSDGLD HNGTGLRVIA QTQDGGTWE QGGVEGKMRG STQTVGLAAK  
 1151 TGENTTAAAT LGMGRSTWSE NSANAKTDSI SLFAGIRHDA GDIGYLGKLF  
 1201 SYGRYKNSIS RSTGADEHAE GSVNGTLMQL GALGGVNVFF AATGDLTVEG  
 20 1251 GLRYDLKQD AFAEKGSALG WSGNSLTEGT LVGLAGLKL QPLSDKAVLF  
 1301 ATAGVERDLN GRDYTVTGGF TGATAATGKT GARNMPHTRL VAGLGADVEF  
 1351 GNGWNLARY SYAGSKQYGN HSGRVGVGYR FLEHHHHHH\*

25

961cL-ORF46.1

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTCTGTAGCGGCGCACTGGCAGCCACAAAC  
 GACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAA  
 GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC  
 30 GACGACTTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAAACAAACGTC  
 GATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAGTTAAACAACCAAGTTAGCAGACACATGATGCCGCTTTAGCA  
 GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAAATATAACGACATTTGCTGAAGAG  
 ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC  
 AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGACAGACGAAGCCGTCAAACCGCCAAATGAAGCCAAACAG  
 35 ACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCC  
 GCTGGCACAGCTAATACTGCAGCCGACAAGCCGAAGCTGCTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT  
 ACGAACAAAGATAATATTGCTAAAAAGCAACAGTGGCGACGTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC  
 AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC  
 GATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAAGAAACCCGCCAAGGCCCTTCGAGAACAAGCC  
 40 GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGAGGAGGATCAGATTTGGCAAACGATTCCTTTT  
 ATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTTCGGCAGCAGGGGGAACTTGCC  
 GATCCGCGCAGGCTATCGGATTGGGAAAAATACAAAGCCATCAGATTGGGCAACCTGATGATTCAACAGGCGGCCATT  
 AAAGGAAATATCGGCTACATTGTCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGACAACCATGCTCA  
 CATTCGATTTCTGATGAAGCCGCTAGTCCCGTTGACGGATTAGCCTTTACCGCATCCATTGGGACGGATACGAACAC  
 45 CATCCGCGCAGCGCTATGACGGGCCACAGGGCGGCGGCTATCCCGCTCCCAAAGGCGGAGGGATATATACAGCTAC  
 GACATAAAAGCGGTTGCCCAAAATATCCGCCTCAACCTGACCGACAACCGCAGCACCAGGACAACGGCTTGCCGACCGT  
 TTCCACAATGCCGGTAGTATGCTGACGCAAGGATAGGCGACGGATTCAAACGCGCCACCCGATACAGCCCCGAGCTG  
 GACAGATCGGGCAATGCCGCGCAAGCCTTCAACGGCATGACAGATATCGTTAAAAACATCATCGGCGCGGACGAGAA  
 ATTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTGCTGTATGCACGGCTTGGGTCTGCTT  
 50 TCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCCGCGCAGGCCATC  
 CGCGATTGGGCGATCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTGAGCAATATCTTTATGGCAGCCATCCCC  
 ATCAAAGGGATTGGAGCTGTTCCGGGAAAAATACGGCTTGGGCGGCATCACGGCACATCCTATCAAGCGGTGCGAGATG  
 GCGCGATCGCATTGCCGAAAGGAAATCCGCGCTCAGCGACAATTTTGGCGATGCGGCATACGCCAAATACCCGTCC  
 CCTTACCATTCCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCTCAACCGTGCCG  
 55 CCGTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGACGGTAAAGGGTTT  
 CCGAATTTTGAGAAGCACGTGAAATATGATACGTAACCTCGAG

55

1 MKHFPSKVL TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQING  
 51 FKAGETIYDI DEDGTITKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN  
 101 KQNVDAKVA AESEIEKLTT KLADTDAALA DTDALDATT NALNLGENI  
 60 151 TTFABETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV  
 201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA  
 251 KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDL  
 301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG  
 351 GSGGGGSDLA NDSFIRQVLD RQHFEPDGKY HLFGRGELA ERSGHIGLKG  
 65 401 IQSHQLGNLM IQQAAIKGNI GYIVRFSHDG HEVHSPFDNH ASHSDSDEAG  
 451 SPVDGFSLYR IHWDGYEHHP ADGYDGPQGG GYPAPKGARD IYSYDIKQVA  
 501 QNIRLNLTDN RSTGQRLADR PHNAGSMLTQ GVGDFGFRAT RYSPELDMSG

551 NAAEAFNGTA DIVKNIIGAA GEIVGAGDAV QGISEGSNIA VMHGLGILLST  
 601 ENKMARINDL ADMAQLKDYA AAAIRDWAVQ NPNAAQGIEA VSNIFMAAIP  
 651 IKGIGAVRGK YGLGGITAHF IKRSQMGAI A LPKGKSAVSD NFADAAYAKY  
 701 PSPYHSRNIR SNLEQRYGKE NITSSTVPPS NGKNVKLADQ RHPKTGVFPD  
 751 KGKFPNFEKH VKYDT\*

**961cL-741**

10 ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTCTGTAGCGGCGCACTGGCAGCCACAAAC  
 GACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA  
 GCTGGAGAGACCATTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC  
 GACGACTTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAACAAACCGTC  
 15 GATGCCAAAGTAAAAGCTGCAGAACTCGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA  
 GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG  
 ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC  
 AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAG  
 ACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCC  
 20 GCTGGCAGAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT  
 ACGAACAAAGATAATATTGCTAAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAAATTTGTC  
 AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC  
 GATACTCCGCTGAACGGTTTGGATAAAACAGTGTGAGACCTGCGCAAGAAACCCGCCAAGGCCCTTCGAGAACAAGCC  
 GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGGGGTGGTGTGCGCCCGACATCGGTGCCGGG  
 25 CTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGG  
 AAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAATTTATGGAACCGGTGACAGCCTCAATACGGGCAAA  
 TTGAAGAACGACAAGGTGAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGACGGGCGAGCTCATTACCTTGGAGAGT  
 GGAGAGTTCCAAGTATACAAACAAGCCATTCCGCCTTAACCGCCTTTCAGACCGAGCAAAATACAAGATTCCGAGCAT  
 TCCGGGAAGATGGTTGCGAAACGCCAGTTTCAAGATCGCGGACATAGCGGGCGAACATACATCTTTTGACAAGCTTCCC  
 GAAGGCGGCGAGGGCGACATATCGCGGGACGGCCTTCGGTTTCAAGCATGCCGGCGGAAAACTGACCTACACCATAGAT  
 30 TTCGCGCCAAGCAGGGAACGGCAAAATCGAACATTTGAAATCGCCAGAACTCAATGTCGACCTGGCCGCCCGCGAT  
 ATCAAGCCGGATGGAAACGCCATGCGGTCTATCAGCGGTTCCGTCTCTTACAACCAAGCCGAGAAAGGCAGTTACTCC  
 CTCGGTATCTTTGGCGGAAAAGCCCAGGAAGTTGCCGGCAGCGCGGAAGTGAAACCGGTAAACGGCATACGCCCATATC  
 GGCCTTGGCGCAAGCAACTCGAGCACCACCACCACCACCTGA

35 1 MKHFPKSVLT TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING  
 51 FKAGETIYDI DEDGTITKRD ATAADVEADD FKGLGLKVV TNLTKTVNE  
 101 KQNVDAKVA AESEIEKLTT KLADTDALA DTDALDATT NALNKLGENI  
 151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV  
 201 KTANEAKQTA BETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA  
 251 KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLD  
 40 301 RLSAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFPQYVNG  
 351 GSGGGVVAAD IGAGLADALT APLDHKDKGL QSLTLDQSVR KNEKLKLAQ  
 401 GAEKTYGNGD SLNTGKLND KVSFRDFIRQ IEVDGQLITL ESGEFQVYKQ  
 451 SHSALTAFQT EQIQDSEHSG KMKVAKRQFRI GDIAGEHTSF DKLPEGGGRAT  
 501 YRGTAFGSDD AGGKLYTID FAAKQNGKI EHLKSPBLNV DLAAADIKPD  
 45 551 GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL  
 601 AAKQLEHHHH HH\*

**961cL-983**

50 ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTCTGTAGCGGCGCACTGGCAGCCACAAAC  
 GACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA  
 GCTGGAGAGACCATTACGACATTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCCGATGTTGAAGCC  
 55 GACGACTTTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAACAAACCGTC  
 GATGCCAAAGTAAAAGCTGCAGAACTCGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA  
 GATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG  
 ACTAAGACAAATATCGTAAAAATTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC  
 AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAG  
 ACGGCCGAAGAAACCAACAAACCGTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCC  
 60 GCTGGCAGAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT  
 ACGAACAAAGATAATATTGCTAAAAAAGCAACAGTGCCGACGTGTACACCAGAGAAGAGTCTGACAGCAAAATTTGTC  
 AGAATTGATGGTCTGAACGCTACTACCGAAAAATTGGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC  
 GATACTCGCCTGAACGGTTTGGATAAAACAGTGTGAGACCTGCGCAAGAAACCCGCCAAGGCCCTTCGAGAACAAGCC  
 GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGCGGAGGGGCACTTCTGCGCCCGACTTCAATGCA  
 65 GCGGTACCGGTATCGGCAGCAACAGCAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCGCGGTATCAAGAAC  
 GAAATGTGCGCAAGACAGAAAGCATGCTGTGCGGTGCGGATGACGTTGCGGTACAGCAGGGATGCCAAATCAAT  
 GCGCGCGCGCGAATCTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT  
 GCAATTGAAGCAGGTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGCGAATCCGTGGCAGCATATCC

TTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAAACTATACGGCGTATATGCGGAAGGAA  
 GCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAGGCCGTTATAGAGACTGAAGCAAAGCCG  
 ACGGATATCCGCCACGTAAGAAATCGGACACATCGATTGGTCTCCCATATTATTGGCGGGCGTTCCGTTGGACGGC  
 5 AGACTCGACGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAAATGATG  
 GTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGCTGGCGTGGCATCGTCAATAACAGTTTGGAAACAACA  
 TCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTTCGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATTCC  
 GGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTACGGCAACCTGTCTTACCACATCCGTAAT  
 AAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT  
 10 GAAAAAGACGCTCAAAAAGGCATTATCACAGTCGACGGCGTAGACCGCAGTGGAGAAAAGTTCAAACGGGAAATGTAT  
 TGAGAACCGGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGCGGAATTACTGCCATGTGGTGCTGTGCGGCACCC  
 TATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTGCGGAACATCCTTTTCCGCACCCATCGTAACC  
 GGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCACGTTGCTGACGACGGCT  
 CAGGACATCGGTGACGTGCGCTGGACAGCAAGTTTCGGCTGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCC  
 GCGTCTTTCCGTTTCGGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCCCTACTCTTCCGTAACGACATT  
 15 TCAGGCACGGCGGCGCTGATCAAAAAGGCGGCAGCCAATGCAACTGCACGGCAACAACACCTATACGGGCAAAACC  
 ATTTATCGAAGGCGGTTCGCTGGTGTGTACGGCAACAACAAATCGGATATGCGCGTCGAAACCAAGGTGCGCTGATT  
 TATAACGGGGCGGCATCCGGCGGCAGCCTGAACAGCAGCAGCAGCATTGCTCTATCTGGCAGATACCGCAATCCGGCGCA  
 AACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAGGTACGCTGTACACACGTTTGGGCAAACTGCTG  
 AAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTGCGCACGGCGCAAGGGGGCAGGCTATCTCAACAGT  
 20 ACCGGACGACGTGTTCCTTCTTGAGTGCCGCCAAAATCGGGCAGGATTATTCTTTCTCACAAACATCGAAACCGAC  
 GCGGCGCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCGACACGCTGTCTATTATGTCCGT  
 CGCGCAATTCGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCGGCTGAAACACCGCGTAGAACAGGGC  
 GCGCAACCTCTGGAACACCTGATGGTCAACTGGATGCCCTCCGAATCATCCGCAACCCGAGACGGTTGAAACTTCGG  
 GCAGCCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTCCGCGCAGCGGCAGCGGTACAGCATGCG  
 25 AATGCCGCGGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCGCCCATGCCGAT  
 ATGCAGGGACGCCGCTGAAAGCGGTATCGGACGGGTGGACCAACGGCACGGGTTCGCGCTCATCGCGCAAAACC  
 CAACAGGACGGTGAACGTTGGGAACAGGGCGGTGTGAAGGCAAAATGCGCGGCAGTACCCAAACCGTCGGCATTGCC  
 CGGAAAACCGGCGAAAATACGACAGCAGCGCCACACTGGGCATGGGACGACGACATGGAGCGAAAACAGTGCAAAT  
 GCAAAAACCGCACAGCATTAGTCTGTTTTCAGGCATACCGGACGATGCGGGCGATATCCGCTATCTCAAAGGCTCTTC  
 30 TCCTACGGACGCTACAAAACAGCATCAGCCGACGACCGGTGCGGACGAACATGCGGAAGGCAGCGTCAACGGCACG  
 CTGATCGACGCTGGGCGCACTGGCGGCTGTCAACGTTCCGTTTTCGCGCAACGGGAGATTGACGGTGAAGGCGGTCTG  
 CGCTACGACCTGCTCAAACAGGATGCATTCCGCCGAAAAGGCAGTGTCTTGGGCTGGAGCGGCAACAGCCTCACTGAA  
 GGCACGCTGGTGGGACTCGCGGCTGTGAAGCTGTGCAACCCCTTGAGCGATAAAGCCGTCCTGTTTGAACGGCGGGC  
 35 GTGGAACCGCAGCTGAACGGACGCGACTACCGGTAACGGCGGCTTTACCGCGCGACTGCGAGCAACCGGCAAGACG  
 GGGGACGCAATATGCCGCACACCCGCTGTTGTCGGCTGGGCGCGGATGTCTGGAATTCGGCAACCGGCTGGAACGGC  
 TTGGCACGTTACAGCTACGCCGTTTCAAACAGTACGGCAACCACAGCGGACGAGTGGCGTAGGCTACCGGTTCTGA  
 CTCGAG

1 MKHFPSKVL TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING  
 40 51 FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKV TNLTKTVNEN  
 101 KQNVDAKVA AESEIEKLTT KLADTDAALA DTDALDATT NALNKLGENI  
 151 TTFAETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTKADEAV  
 201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA  
 45 251 KVTDIKADIA TNKNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDI  
 301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFPQYNVG  
 351 GSGGGTSAP DFNAGGTGIG SNSRATTAKS AAVSYAGIKN EMCKDRSMILC  
 401 AGRDDVAVD RDAKINAPPP NLHTGDFPNP NDAYKNLINL KPAIEAGYTG  
 451 RGVEVGIVDT GESVGSISFP ELYGRKEHGY NENYKNYTAY MRKEAPEDGG  
 501 GKDIEASFDD EAVIETBAKP TDIRHVKEIG HIDLVSHIIG GRSVDGRPAG  
 50 551 GIAPDATHI MNTNDETKNE MMVAIRNAW VKLGERGVRI VNNSFGTTSR  
 601 AGTADLFQIA NSEEQYRQAL LDYSGGDKTD EGIRLMQSD YGNLSYHIRN  
 651 KNMLFIFSTG NDAQAPNTY ALLPFYEKDA QKGIITVAGV DRSGEKFKRE  
 701 MYGEPGTEPL EYGSNHCGIT AMWCLAPYE ASVRFRTRNP IQIAGTSFSA  
 751 PIVTGTAALL LQKYPWMSND NLRTLLTTA QDIGAVGVS KFGWGLLDAG  
 55 801 KAMNGPASFP FGDFADTKG TSDIAYSFRN DISGTGGLIK KGGSQLQLHG  
 851 NNTYTGTII EGGSLVLYGN NKSDMRVETK GALIYNGAAS GGSLSNDGIV  
 901 YLADTDQSGA NETVHIKSL QLDGKGTLYT RLKLLKVDG TAIIGKLYM  
 951 SARGKGAGYL NSTGRRVPFL SAKIGQDYS FFTNIETDGG LLASLDSVEK  
 1001 TAGSEGDTLS YYVRRGNAAR TASAAHSAP AGLKHAVEQG GSNLENLMVE  
 60 1051 LDASESATP ETVETAAADR TDMPGIRPYG ATFRAAAVQ HANAADGVRI  
 1101 FNSLAATVYA DSTAAHADMQ GRRLKAVSDG LDHNGTGLRV IAQTQDGGT  
 1151 WEQGGVEGKM RGSTQTVGIA AKTGENTTAA ATLGMRSTW SENSANAKTD  
 1201 SISLFAIRH DAGDIGYKLG LPSYGRYKNS ISRSTGADEH AEGSVNGTLM  
 1251 QLALGGMNV PFAATGDLTV EGGLEYDLLK QDAFAEKSA LGWSGNSLTE  
 65 1301 GTLVGLAGLK LSQPLSDKAV LFATAGVERD LNGRDYTVTG GFTGATAATG  
 1351 KTGARNMPHT RLVAGLGADV EFGNGWNGLA RYSYAGSKQY GNHSGRVGVG  
 1401 YRF\*

It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention. For instance, the use of proteins from other strains is envisaged [e.g. see WO00/66741 for polymorphic sequences for ORF4, ORF40, ORF46, 225, 235, 287, 519, 726, 919 and 953].

5

## EXPERIMENTAL DETAILS

### *Cloning strategy and oligonucleotide design*

Genes coding for antigens of interest were amplified by PCR, using oligonucleotides designed on the basis of the genomic sequence of *N. meningitidis* B MC58. Genomic DNA from strain 2996 was always used as a template in PCR reactions, unless otherwise specified, and the amplified fragments were cloned in the expression vector pET21b+ (Novagen) to express the protein as C-terminal His-tagged product, or in pET-24b+(Novagen) to express the protein in 'untagged' form (e.g. ΔG 287K).

Where a protein was expressed without a fusion partner and with its own leader peptide (if present), amplification of the open reading frame (ATG to STOP codons) was performed.

Where a protein was expressed in 'untagged' form, the leader peptide was omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

The melting temperature of the primers used in PCR depended on the number and type of hybridising nucleotides in the whole primer, and was determined using the formulae:

$$T_{m1} = 4 (G+C) + 2 (A+T) \quad (\text{tail excluded})$$

$$T_{m2} = 64.9 + 0.41 (\% \text{ GC}) - 600/N \quad (\text{whole primer})$$

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-60°C for the hybridising region alone.

Oligonucleotides were synthesised using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH<sub>4</sub>OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in water.

		Sequences	Restriction site
fu (961 )-	Fwd	CGCGGATCC -GGAGGGGGTGGTGTCTG	BamHI

741(MC58)-His			
	Rev	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
fu (961 )-983-His	Fwd	CGCGGATCC - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
fu (961)- Orf46.1-His	Fwd	CGCGGATCCGGTGGTGGTGGT-TCAGATTGGCAAACGATTC	BamHI
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI
fu (961 c-L)-741(MC58)	Fwd	CGCGGATCC -GGAGGGGGTGGTGTCTG	BamHI
	Rev	CCCGCTCGAG-TTATTGCTTGGCGGCAAG	XhoI
fu (961c-L )-983	Fwd	CGCGGATCC - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCGCTCGAG-TCAGAACCGGTAGCCTAC	XhoI
fu (961c-L)-Orf46.1	Fwd	CGCGGATCCGGTGGTGGTGGT-TCAGATTGGCAAACGATTC	BamHI
	Rev	CCCGCTCGAG-TTACGTATCATATTTACGTGC	XhoI
fu-(ΔG287)-919-His	Fwd	CGCGGATCCGGTGGTGGTGGT-CAAAGCAAGAGCATCCAAACC	BamHI
	Rev	CCCAAGCTT-TTCGGGCGGTATTCGGGCTTC	HindIII
fu-(ΔG287)-953-His	Fwd	CGCGGATCCGGTGGTGGTGGT-GCCACCTACAAAGTGGAC	BamHI
	Rev	GCCCAAGCTT-TTGTTGGCTGCCTCGAT	HindIII
fu-(ΔG287)-961-His	Fwd	CGCGGATCCGGTGGTGGTGGT-ACAAGCGACGACG	BamHI
	Rev	GCCCAAGCTT-CCACTCGTAATTGACGCC	HindIII
fu-(ΔG287)-Orf46.1-His	Fwd	CGCGGATCCGGTGGTGGTGGT-TCAGATTGGCAAACGATTC	BamHI
	Rev	CCCAAGCTT-CGTATCATATTTACGTGC	HindIII
fu-(ΔG287-919)-Orf46.1-His	Fwd	CCCAAGCTTGGTGGTGGTGGTGGT-TCAGATTGGCAAACGATTC	HindIII
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI
fu-(ΔG287-Orf46.1)-919-His	Fwd	CCCAAGCTTGGTGGTGGTGGTGGT-CAAAGCAAGAGCATCCAAACC	HindIII
	Rev	CCCGCTCGAG-CGGGCGGTATTCGGGCTT	XhoI
fu ΔG287(394.98)-...	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	NheI
	Rev	CGGGGATCC-ATCCTGCTCTTTTTTGCCGG	BamHI
fu Orf1-(Orf46.1)-His	Fwd	CGCGGATCCGCTAGC-GGACACACTTATTTGCGCATC	NheI
	Rev	CGCGGATCC-CCAGCGGTAGCCTAATTTGAT	
fu (Orf1)-Orf46.1-His	Fwd	CGCGGATCCGGTGGTGGTGGT-TCAGATTGGCAAACGATTC	BamHI
	Rev	CCCAAGCTT-CGTATCATATTTACGTGC	HindIII
fu (919)-Orf46.1-His	Fwd1	GCGGCGTTCGACGGTGGCGGAGGCACTGGATCCTCAG	SalI
	Fwd2	GGAGGCACTGGATCCTCAGATTGGCAAACGATTC	
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI
Fu (orf46)-287-His	Fwd	CGGGGATCCGGGGGCGGCGGTGGCG	BamHI
	Rev	CCCAAGCTTATCCTGCTCTTTTTTGCCGGC	HindIII
Fu (orf46)-919-His	Fwd	CGCGGATCCGGTGGTGGTGGTCAAAGCAAGAGCATCCA AACC	BamHI
	Rev	CCCAAGCTTCGGGCGGTATTCGGGCTTC	HindIII

Fu (orf46-919)- 287-His	Fwd	CCCCAAGCTTGGGGGCGGCGGTGGCG	HindIII
	Rev	CCCGCTCGAGATCCTGCTCTTTTGGCCGGC	XhoI
Fu (orf46-287)- 919-His	Fwd	CCCAAGCTTGGTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC	HindIII
	Rev	CCCGCTCGAGCGGGCGGTATTCGGGCTT	XhoI
(ΔG741)-961c-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAGGTTG	XhoI
(ΔG741)-961-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG741)-983-His	Fwd	GCGGCCTCGAG- GGATCCGGCGGAGGCGGCACTTCTGCG	XhoI
	Rev	CCCGCTCGAG-GAACC GG TAGCCTACG	XhoI
(ΔG741)-orf46.1- His	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
	Fwd2	GCGGCGTTCGACGGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI
(ΔG983)- 741(MC58)-His	Fwd	GCGGCCTCGAG-GGATCCGGAGGGGGTGGTGTGCGCC	XhoI
	Rev	CCCGCTCGAG-TTGCTTGGCGGCAAG	XhoI
(ΔG983)-961c-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAGGTTG	XhoI
(ΔG983)-961-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG983)-Orf46.1- His	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
	Fwd2	GCGGCGTTCGACGGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI

\* This primer was used as a Reverse primer for all the C terminal fusions of 287 to the His-tag.

<sup>§</sup> Forward primers used in combination with the 287-His Reverse primer.

NB – All PCR reactions use strain 2996 unless otherwise specified (e.g. strain MC58)

In all constructs starting with an ATG not followed by a unique *NheI* site, the ATG codon is  
 5 part of the *NdeI* site used for cloning. The constructs made using *NheI* as a cloning site at the  
 5' end (e.g. all those containing 287 at the N-terminus) have two additional codons (GCT  
 AGC) fused to the coding sequence of the antigen.

### ***Preparation of chromosomal DNA templates***

*N.meningitidis* strains 2996, MC58, 394.98, 1000 and BZ232 (and others) were grown to  
 10 exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in  
 5ml buffer (20% w/v sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes  
 incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1%  
 Na-Sarkosyl, 50μg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two

phenol extractions (equilibrated to pH 8) and one  $\text{CHCl}_3$ /isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation. The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The  
5 DNA concentration was measured by reading  $\text{OD}_{260}$ .

### ***PCR Amplification***

The standard PCR protocol was as follows: 200ng of genomic DNA from 2996, MC581000, or BZ232 strains or 10ng of plasmid DNA preparation of recombinant clones were used as template in the presence of 40 $\mu\text{M}$  of each oligonucleotide primer, 400-800  $\mu\text{M}$  dNTPs solution, 1x PCR buffer (including 1.5mM  $\text{MgCl}_2$ ), 2.5 units *TaqI* DNA polymerase (using  
10 Perkin-Elmer AmpliTaq, Boehringer Mannheim Expand<sup>TM</sup> Long Template).

After a preliminary 3 minute incubation of the whole mix at 95°C, each sample underwent a two-step amplification: the first 5 cycles were performed using the hybridisation temperature that excluded the restriction enzyme tail of the primer ( $T_{m1}$ ). This was followed by 30 cycles  
15 according to the hybridisation temperature calculated for the whole length oligos ( $T_{m2}$ ). Elongation times, performed at 68°C or 72°C, varied according to the length of the Orf to be amplified. In the case of Orf1 the elongation time, starting from 3 minutes, was increased by 15 seconds each cycle. The cycles were completed with a 10 minute extension step at 72°C.

The amplified DNA was either loaded directly on a 1% agarose gel. The DNA fragment  
20 corresponding to the band of correct size was purified from the gel using the Qiagen Gel Extraction Kit, following the manufacturer's protocol.

### ***Digestion of PCR fragments and of the cloning vectors***

The purified DNA corresponding to the amplified fragment was digested with the appropriate restriction enzymes for cloning into pET-21b+, pET22b+ or pET-24b+. Digested  
25 fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted with either  $\text{H}_2\text{O}$  or 10mM Tris, pH 8.5. Plasmid vectors were digested with the appropriate restriction enzymes, loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit.

### ***Cloning***

The fragments corresponding to each gene, previously digested and purified, were ligated into pET21b+, pET22b+ or pET-24b+. A molar ratio of 3:1 fragment/vector was used with T4 DNA ligase in the ligation buffer supplied by the manufacturer.

- 5 Recombinant plasmid was transformed into competent *E.coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice, then at 37°C for 3 minutes. This was followed by the addition of 800µl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200µl of the supernatant and plated onto LB ampicillin (100mg/ml ) agar.
- 10 Screening for recombinant clones was performed by growing randomly selected colonies overnight at 37°C in 4.0ml of LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel
- 15 (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

### ***Expression***

- After cloning each gene into the expression vector, recombinant plasmids were transformed into *E.coli* strains suitable for expression of the recombinant protein. 1µl of each construct
- 20 was used to transform *E.coli* BL21-DE3 as described above. Single recombinant colonies were inoculated into 2ml LB+Amp (100µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20ml of LB+Amp (100µg/ml) in 100ml flasks, to give an OD<sub>600</sub> between 0.1 and 0.2. The flasks were incubated at 30°C or at 37°C in a gyratory water bath shaker until OD<sub>600</sub> indicated exponential growth suitable for induction of expression (0.4-0.8 OD). Protein
  - 25 expression was induced by addition of 1.0mM IPTG. After 3 hours incubation at 30°C or 37°C the OD<sub>600</sub> was measured and expression examined. 1.0ml of each sample was centrifuged in a microfuge, the pellet resuspended in PBS and analysed by SDS-PAGE and Coomassie Blue staining.

### ***Purification of His-tagged proteins***

- 30 Various forms of 287 were cloned from strains 2996 and MC58. They were constructed with a C-terminus His-tagged fusion and included a mature form (aa 18-427), constructs with



deletions ( $\Delta 1$ ,  $\Delta 2$ ,  $\Delta 3$  and  $\Delta 4$ ) and clones composed of either B or C domains. For each clone purified as a His-fusion, a single colony was streaked and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking.

5 The overnight culture was diluted 1:30 into 1.0 L LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (30 or 37°C) until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000g for 15 min at 4°C. The bacterial pellet was resuspended in 7.5 ml of

10 either (i) cold buffer A (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (10mM Tris-HCl, 100 mM phosphate buffer, pH 8.8 and, optionally, 8M urea) for insoluble proteins. Proteins purified in a soluble form included 287-His,  $\Delta 1$ ,  $\Delta 2$ ,  $\Delta 3$  and  $\Delta 4$ 287-His,  $\Delta 4$ 287MC58-His, 287c-His and 287cMC58-His. Protein 287bMC58-His was insoluble and purified accordingly. Cells were disrupted by

15 sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6 M guanidine hydrochloride, 100 mM phosphate buffer, 10 mM Tris- HCl, pH 7.5 and treated with 10 passes of a Dounce homogenizer. The homogenate was centrifuged at 13000g for 30 min and the supernatant retained. Supernatants for both soluble and insoluble

20 preparations were mixed with 150µl Ni<sup>2+</sup>-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to the manufacturer's protocol. The batch-wise preparation was centrifuged at 700g for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml

25 buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD<sub>280</sub> of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (10mM Tris-HCl, 100mM phosphate buffer, pH 6.3 and, optionally, 8M urea) until OD<sub>280</sub> of the flow-through reached 0.02-0.01. The His-fusion

30 protein was eluted by addition of 700µl of either (i) cold elution buffer A (300 mM NaCl, 50mM phosphate buffer, 250 mM imidazole, pH 8.0) or (ii) elution buffer B (10 mM Tris-HCl, 100 mM phosphate buffer, pH 4.5 and, optionally, 8M urea) and fractions

collected until the OD<sub>280</sub> indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analysed by SDS-PAGE. Protein concentrations were estimated using the Bradford assay.

#### ***Renaturation of denatured His-fusion proteins.***

- 5 Denaturation was required to solubilize 287bMC8, so a renaturation step was employed prior to immunisation. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10% v/v. The proteins were diluted to 200 µg/ml using dialysis buffer I (10% v/v glycerol, 0.5M arginine, 50 mM phosphate buffer, 5.0 mM reduced glutathione, 0.5 mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for  
10 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% v/v glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C. Protein concentration was estimated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

#### ***Immunization***

- 15 Balb/C mice were immunized with antigens on days 0, 21 and 35 and sera analyzed at day 49.

#### ***Sera analysis – ELISA***

- The acapsulated MenB M7 and the capsulated strains were plated on chocolate agar plates and incubated overnight at 37°C with 5% CO<sub>2</sub>. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco)  
20 containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.4-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and bacteria were washed twice with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells  
25 were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN<sub>3</sub> in PBS) were added to  
30 each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells

were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-phenildiamine and 10µl of H<sub>2</sub>O<sub>2</sub>) were added to each well and the plates were left at room temperature for 20 minutes. 100µl 12.5% H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD<sub>490</sub> value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD<sub>490</sub> of 0.4 was higher than 1:400.

#### ***Sera analysis – FACS Scan bacteria binding assay***

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C with 5% CO<sub>2</sub>. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN<sub>3</sub>) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD<sub>620</sub> of 0.05. 100µl bacterial cells were added to each well of a Costar 96 well plate. 100µl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200µl/well of blocking buffer in each well. 100µl of R-Phicoerytrin conjugated F(ab)<sub>2</sub> goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200µl/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200µl/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

#### ***Sera analysis – bactericidal assay***

*N. meningitidis* strain 2996 was grown overnight at 37°C on chocolate agar plates (starting from a frozen stock) with 5% CO<sub>2</sub>. Colonies were collected and used to inoculate 7ml Mueller-Hinton broth, containing 0.25% glucose to reach an OD<sub>620</sub> of 0.05-0.08. The culture was incubated for approximately 1.5 hours at 37 degrees with shaking until the OD<sub>620</sub> reached the value of 0.23-0.24. Bacteria were diluted in 50mM Phosphate buffer pH 7.2 containing 10mM MgCl<sub>2</sub>, 10mM CaCl<sub>2</sub> and 0.5% (w/v) BSA (assay buffer) at the working dilution of 10<sup>5</sup> CFU/ml. The total volume of the final reaction mixture was 50 µl with 25 µl

of serial two fold dilution of test serum, 12.5 µl of bacteria at the working dilution, 12.5 µl of baby rabbit complement (final concentration 25% ).

Controls included bacteria incubated with complement serum, immune sera incubated with bacteria and with complement inactivated by heating at 56°C for 30'. Immediately after the addition of the baby rabbit complement, 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 0). The 96-wells plate was incubated for 1 hour at 37°C with rotation. 7µl of each sample were plated on Mueller-Hinton agar plates as spots, whereas 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 1). Agar plates were incubated for 18 hours at 37 degrees and the colonies corresponding to time 0 and time 1 were counted.

#### ***Sera analysis – western blots***

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3% Tris base, 1.44% glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

The OMVs were prepared as follows: *N. meningitidis* strain 2996 was grown overnight at 37 degrees with 5% CO<sub>2</sub> on 5 GC plates, harvested with a loop and resuspended in 10 ml of 20mM Tris-HCl pH 7.5, 2 mM EDTA. Heat inactivation was performed at 56°C for 45 minutes and the bacteria disrupted by sonication for 5 minutes on ice (50% duty cycle, 50% output, Branson sonifier 3 mm microtip). Unbroken cells were removed by centrifugation at 5000g for 10 minutes, the supernatant containing the total cell envelope fraction recovered and further centrifuged overnight at 50000g at the temperature of 4°C. The pellet containing the membranes was resuspended in 2% sarkosyl, 20mM Tris-HCl pH 7.5, 2 mM EDTA and incubated at room temperature for 20 minutes to solubilise the inner membranes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, the supernatant was further centrifuged at 50000g for 3 hours. The pellet, containing the outer membranes

was washed in PBS and resuspended in the same buffer. Protein concentration was measured by the D.C. Bio-Rad Protein assay (Modified Lowry method), using BSA as a standard.

Total cell extracts were prepared as follows: *N. meningitidis* strain 2996 was grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl.

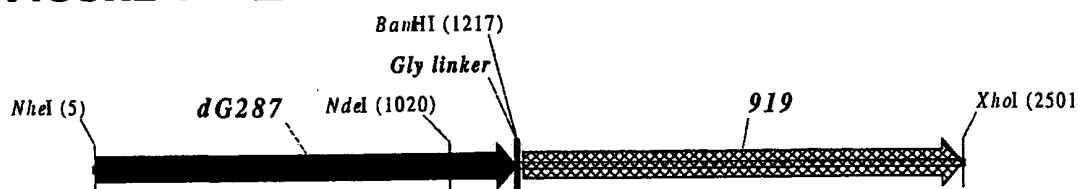
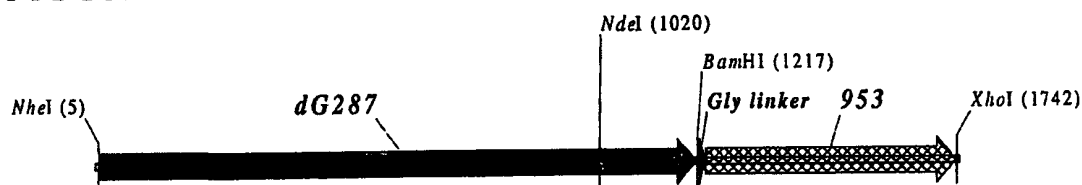
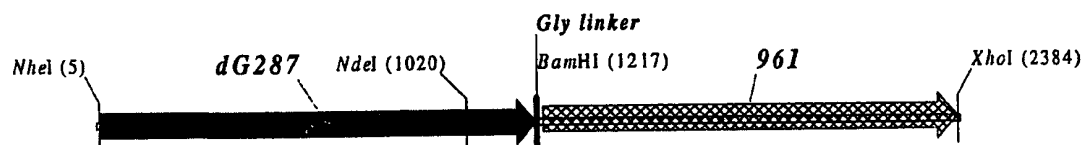
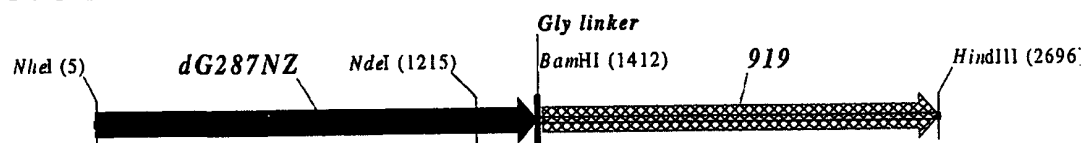
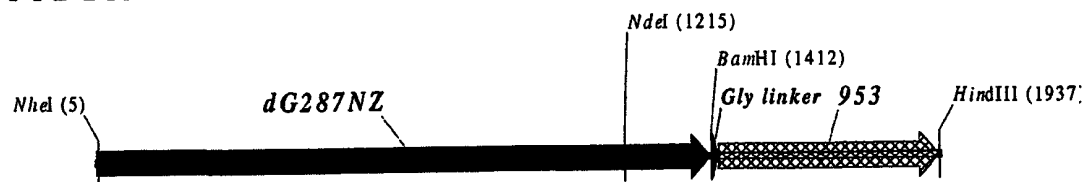
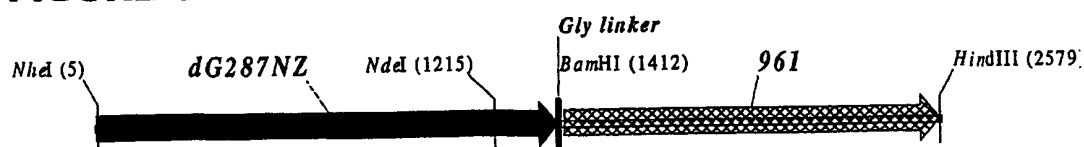
- 5 Heat inactivation was performed at 56°C for 30 minutes.

**CLAIMS**

1. A method for the simultaneous heterologous expression of two or more proteins of the invention, in which (a) two or more proteins of the invention are fused.
2. The method of claim 23, in which the two or more proteins are: (a) 919 and 287; (b) 953  
5 and 287; (c) 287 and ORF46.1; (d) ORF1 and ORF46.1; (e) 919 and ORF46.1; (f)  
ORF46.1, 287 and 919; (g) 919 and 519; and (h) ORF97 and 225.
3. The method of claim 24, in which 287 is at the C-terminal end of protein (a), (b) or (c).
4. The method of any preceding claim, in which the expression is in an *E.coli* host.
5. A protein expressed by the method of any preceding claim.
- 10 6. A hybrid protein of formula  $\text{NH}_2\text{-A-B-COOH}$ , wherein A and B are different Neisserial proteins.
7. The protein of claim 6, wherein A and B are each selected from orf1, orf4, orf25, orf40, orf46, orf83, 233, 287, 2921, 564, 687, 741, 907, 919, 953, 961 and 983.
8. The protein of claim 7, wherein A and B are each selected from ORF46.1, 287, 741, 919,  
15 953, 961 and 983.
9. The protein of claim 8, wherein at least one of said ORF46.1, 287, 741, 919, 953, 961 and 983 is used in essentially full-length form
10. The protein of claim 8 or claim 9, wherein at least one of said ORF46.1, 287, 741, 919, 953, 961 and 983 has a deletion.
- 20 11. The protein of claim 10, wherein A and/or B has a poly-glycine deletion (' $\Delta\text{G}$ ').
12. The protein of claim 11, wherein A and/or B is  $\Delta\text{G-287}$ ,  $\Delta\text{GTbp2}$ ,  $\Delta\text{G741}$ , or  $\Delta\text{G983}$ .
13. The protein of claim 10, wherein A and/or B is a truncated protein.
14. The protein of claim 13, wherein A and/or B is  $\Delta\text{1-287}$ ,  $\Delta\text{2-287}$ ,  $\Delta\text{3-287}$  or  $\Delta\text{4-287}$ .
15. The protein of claim 10, wherein a domain of A and/or B is deleted.

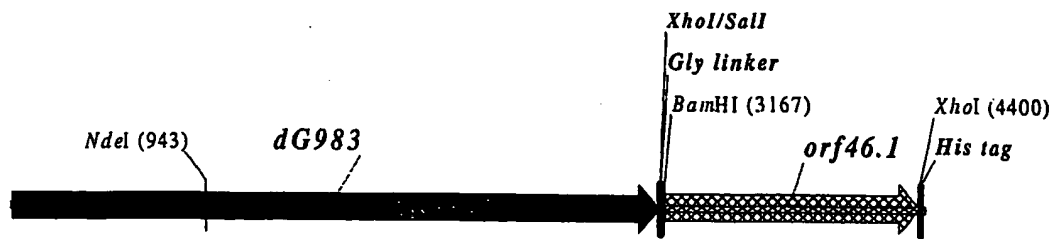
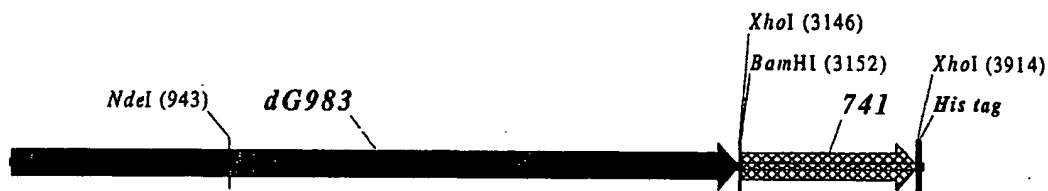
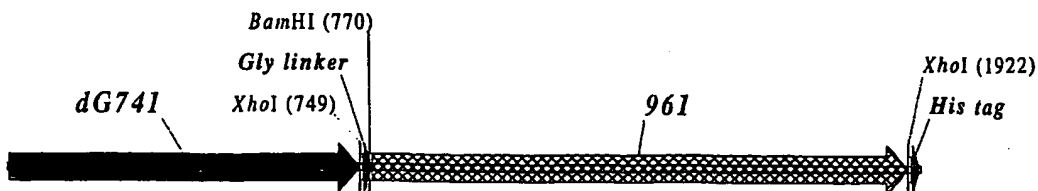
16. The protein of claim 15, wherein A and/or B is 287B, 287C, 287BC, ORF461-433, ORF46433-608, ORF46, or 961c.
17. The protein of claim 6, wherein A and B are: (a) 919 and 287; (b) 953 and 287; (c) 287 and ORF46.1; (d) ORF1 and ORF46.1; (e) 919 and ORF46.1; (f) ORF46.1 and 919; (g) 919 and 519; or (h) ORF97 and 225.
18. The protein of claim 17, wherein the protein is  $\Delta$ G287-919,  $\Delta$ G287-953,  $\Delta$ G287-961,  $\Delta$ G983-ORF46.1,  $\Delta$ G983-741,  $\Delta$ G983-961,  $\Delta$ G983-961C,  $\Delta$ G741-961,  $\Delta$ G741-961C,  $\Delta$ G741-983,  $\Delta$ G741-ORF46.1, ORF46.1-741, ORF46.1-961, ORF46.1-961C, 961-ORF46.1, 961-741, 961-983, 961C-ORF46.1, 961C-741, 961C-983, 961CL-ORF46.1, 961CL-741, or 961CL-983.
19. The protein of claim 8, wherein A or B is 287.
20. The protein of claim 19, wherein B is 287
21. The protein of claim 19, wherein A is  $\Delta$ G-287
22. The protein of claim 21, wherein B is ORF46, 919, 953 or 961.
23. The protein of any one of claims 19 to 22, wherein 287 is from strain 2996 or 394/98.
24. The protein of claim 8, wherein A is 961.
25. The protein of any one of claims 6 to 24, wherein A and B are from the same strain.
26. The protein of any one of claims 6 to 24, wherein A and B are joined directly
27. The protein of any one of claims 6 to 24, wherein A and B are joined via a linker peptide.
28. The protein of claim 27, wherein the linker peptide is a poly-glycine linker, with the proviso that B is not a  $\Delta$ G protein.

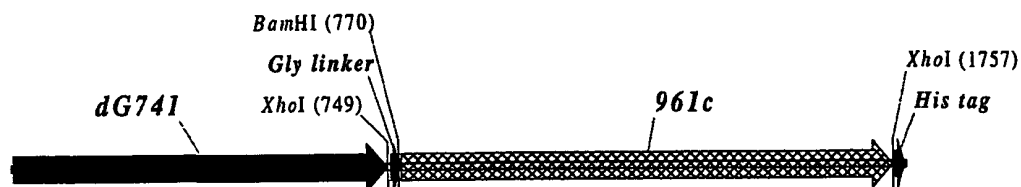
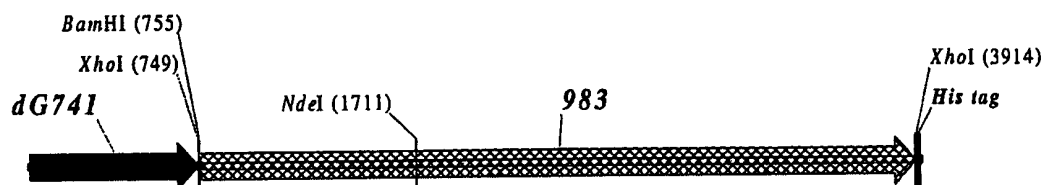
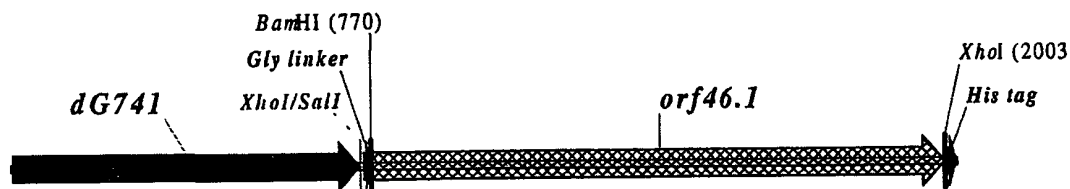
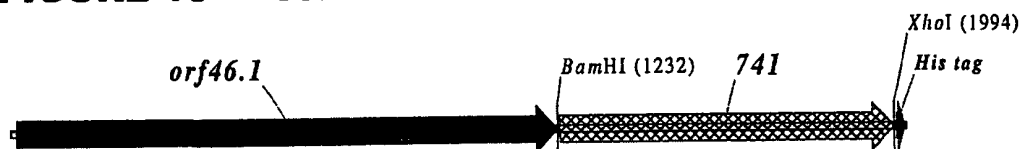
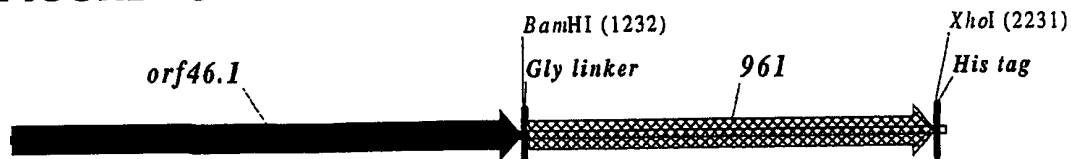
1/5

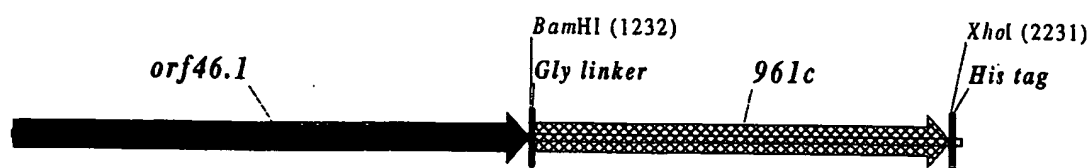
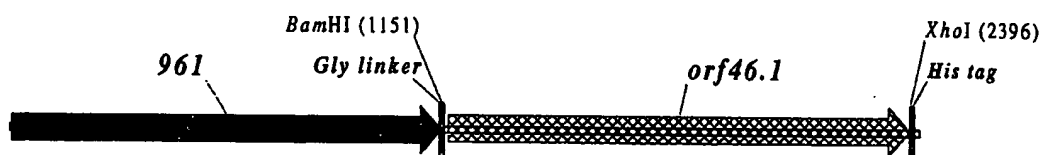
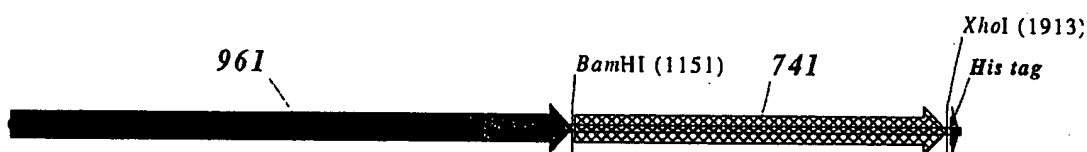
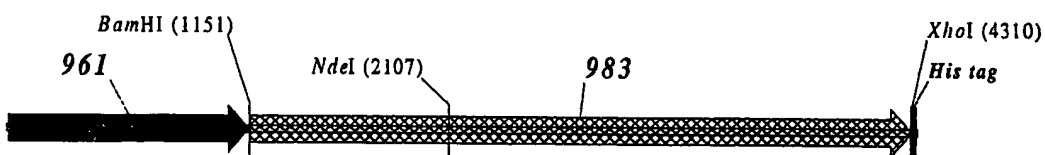
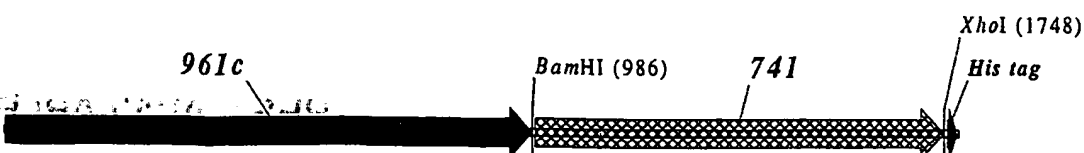
**FIGURE 1 — ΔG287—919****FIGURE 2 — ΔG287—953****FIGURE 3 — ΔG287—961****FIGURE 4 — ΔG287NZ—919****FIGURE 5 — ΔG287NZ—953****FIGURE 6 — ΔG287NZ—961**

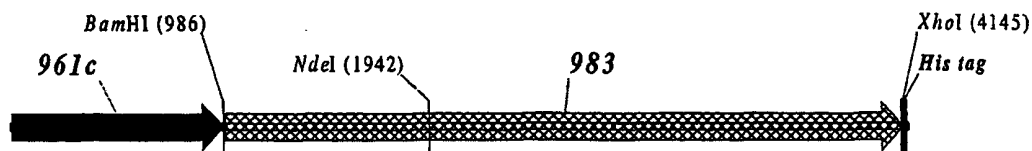
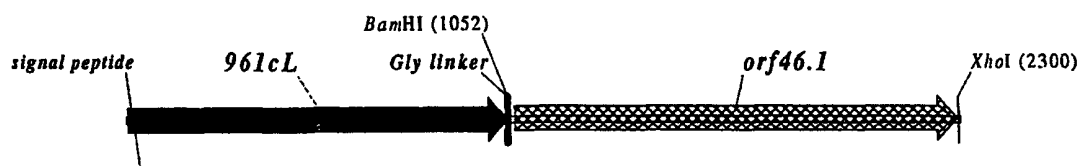
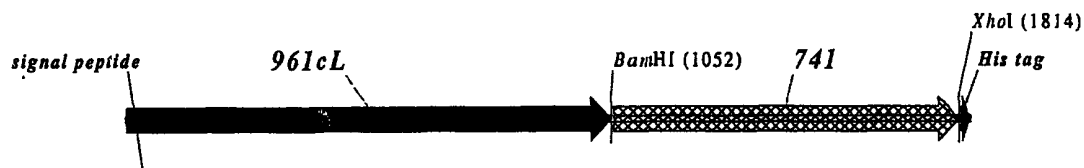
BEST AVAILABLE COPY



**FIGURE 7 — ΔG983-ORF46.1****FIGURE 8 — ΔG983-741****FIGURE 9 — ΔG983-961****FIGURE 10 — ΔG983-961c****FIGURE 11 — ΔG741-961**

**FIGURE 12 — ΔG741-961c****FIGURE 13 — ΔG741-983****FIGURE 14 — ΔG741-ORF46.1****FIGURE 15 — ORF46.1-741****FIGURE 16 — ORF46.1-961**

**FIGURE 17 — ORF46.1—961c****FIGURE 18 — 961-ORF46.1****FIGURE 19 — 961-741****FIGURE 20 — 961-983****FIGURE 21 — 961c-ORF46.1****FIGURE 22 — 961c-741**

**FIGURE 23 — 961c-983****FIGURE 24 — 961cL-ORF46.1****FIGURE 25 — 961cL-741****FIGURE 26 — 961cL-983**